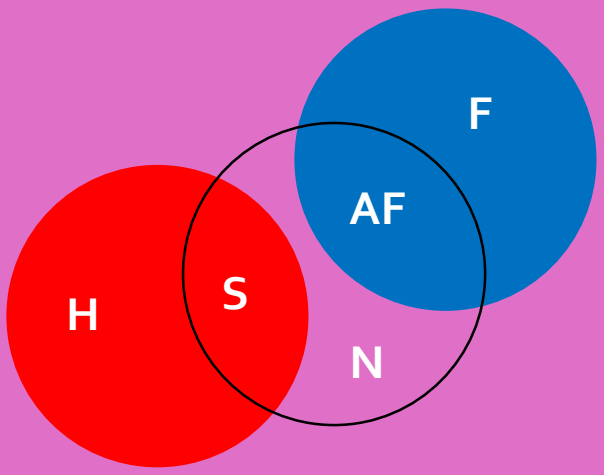
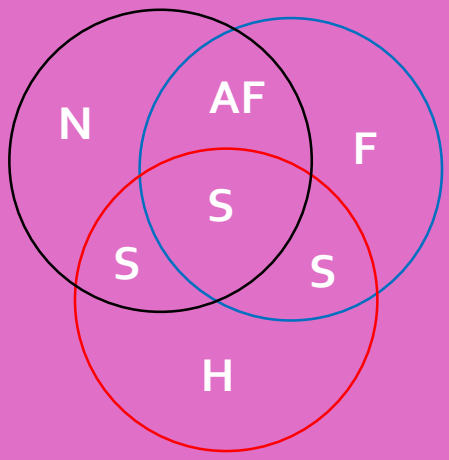
Florentin Smarandache & Jean Dezert
Editors

Advances and Applications of DSmT for Information Fusion

Collected Works. Volume 4



Venn diagram for the STANAG allegiances.



Venn diagram for a possible combat situation.

American Research Press (ARP)
2015

Florentin Smarandache & Jean Dezert (Editors)

Advances and Applications of DSmT for Information Fusion

Collected Works. Volume 4

In memoriam of our colleagues and friends

Jean-Pierre Le Cadre (1953-2009)
Pierre Valin (1949-2014)
Darko Mušicki (1957-2014)

American Research Press (ARP)

2015

Freely download this book from http://fs.gallup.unm.edu/DSmT-book4.pdf.

Download books from Digital Library of Science at:
http://fs.gallup.unm.edu//eBooks-otherformats.html.

This book is also in Amazon Search Inside the Book, Amazon Kindle, and in Google Book Search.

Download Acrobat Reader 11.0.06 (recommended version) for best exploration of this ebook:
http://www.adobe.com/support/downloads/detail.jsp?ftpID=5718 (for Macintosh platform)
http://www.adobe.com/support/downloads/detail.jsp?ftpID=5715 (for Windows platform).

ISSN-10: 1-59973-073-1

ISBN-13: 978-1-59973-324-1

The ISSN 1938-5544 has been assigned by The Library of Congress, Washington DC, USA, for this book series.

© 2015 by the Editors, the Authors for their articles, and the Publishing House.

This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying the content (or some part(s) of it) are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder, and without giving the exact original reference or source.

Table of Contents

Foreword / 7

Part 1: Theoretical advances on DSmT / 9

Non Bayesian conditioning and deconditioning / 11
by Dezert J., Smarandache F.

Multi-criteria decision making based on DSmT/AHP / 17
by Dezert J, Tacnet J.-M., Batton-Hubert M., Smarandache F.

Sequential adaptive combination of unreliable sources of evidence / 23
by Liu Z., Pan Q., Cheng Y., Dezert J.

Fusion of sources of evidence with different importances and reliabilities / 29
by Smarandache F., Dezert J., Tacnet J.-M.

Is Entropy Enough to Evaluate the Probability Transformation Approach of Belief Function? / 37
by Han D., Dezert J., Han C., Yang Y.

Algebraic Generalization of Venn Diagram / 45
by Smarandache F.

Importance of Sources using the Repeated Fusion Method and the Proportional Conflict Redistribution Rules #5 and #6 / 47
by Smarandache F., Dezert J.

Evidence Supporting Measure of Similarity for Reducing the Complexity in Information Fusion / 53
by Li X., Dezert J., Smarandache F., Huang X.

Contradiction Measures and Specificity Degrees of Basic Belief Assignments / 87
by Smarandache F., Martin A., Osswald C.

Evidential Reasoning for Multi-Criteria Analysis based on DSmT-AHP / 95
by Dezert J., Tacnet J.-M.

Cautious OWA and Evidential Reasoning for Decision Making under Uncertainty / 101
by Tacnet J.-M., Dezert J.

Sigmoidal Model for Belief Function-based Electre Tri method / 109
by Dezert J., Tacnet J.-M.

Hierarchical proportional redistribution for bba approximation / 117
by Dezert J., Han D., Liu Z., Tacnet J.-M.

Hierarchical DSmP transformation for decision-making under uncertainty / 125
by Dezert J., Han D., Liu Z., Tacnet J.-M.

Neutrosophic Masses & Indeterminate Models / 133
by Smarandache F.

Extended PCR Rules for Dynamic Frames / 139
by Smarandache F., Dezert J.

A Fuzzy-Cautious OWA Approach with Evidential Reasoning / 147
by Han D., Dezert J., Tacnet J.-M., Han C.

Soft ELECTRE TRI outranking method based on belief functions / 155
by Dezert J., Tacnet J.-M.

On The Validity of Dempster-Shafer Theory / 163
by Dezert J., Wang P., Tchamova A.

Hierarchical Proportional Redistribution principle
for uncertainty reduction and bba approximation / 169
by Dezert J., Han D., Liu Z., Tacnet J.-M.

On the Behavior of Dempster's Rule of Combination and the Foundations
of Dempster-Shafer Theory / 177
by Tchamova A., Dezert J.

Comparative Study of Contradiction Measures in the Theory of Belief Functions / 183
by Smarandache F., Han D., Martin A.

Why Dempster's rule doesn't behave as Bayes rule with informative priors / 189
by Dezert J., Tchamova A., Han D., Tacnet J.-M.

Why Dempster's fusion rule is not a generalization of Bayes fusion rule / 195
by Dezert J., Tchamova A., Han D., Tacnet J.-M.

On the consistency of PCR6 with the averaging rule
and its application to probability estimation / 203
by Smarandache F., Dezert J.

Examples where the conjunctive and Dempster's rules are insensitive / 211
by Smarandache F., Dezert J., Kroumov V.

URREF Self-Confidence in Information Fusion Trust / 215
by Blasch E., Jøsang A., Dezert J., Costa P., Jousselme A.-L.

Evaluations on Evidence Combination Rules in Terms of Statistical Sensitivity and Divergence / 223
by Han D., Dezert J., Yang Y.

A new self-adaptive fusion algorithm based on DST and DSmT / 231
by Yu X.H., Zhou Q.-J., Li Y.-L., An J., Liu Z.-C.

On the quality of optimal assignment for data association / 239
by Dezert J., Benameur K.

Comparison of Identity Fusion Algorithms Using Estimations of Confusion Matrices / 249
by Golino G., Graziano A., Farina A., Mellano W., Ciaramaglia F.

Reliability and Importance Discounting of Neutrosophic Masses / 257
by Smarandache F.

Part 2: Applications of DSmT / 269

DSm theory for fusing highly conflicting ESM reports / 271
by Valin P., Djiknavorian P., Grenier D.

An application of DSmT in ontology-based fusion systems / 279
by Krenc K., Kawalec A.

GMTI and IMINT data fusion for multiple target tracking and classification / 287
by Pannetier B., Dezert J.

Threat assessment of a possible Vehicle-Born Improvised Explosive Device using DSmT / 295
by Dezert J., Smarandache F.

A PCR-BIMM filter for maneuvering target tracking / 321
by Dezert J., Pannetier B.

Implementation of Approximations of Belief Functions for Fusion of ESM Reports within the DSm Framework / 329
by Djiknavorian P., Valin P., Grenier D.

Multiple ground target tracking and classification with DSmT / 337
by Pannetier B., Dezert J.

Edge Detection in Color Images Based on DSmT / 343
by Dezert J., Liu Z., Mercier G.

DSmT Applied to Seismic and Acoustic Sensor Fusion / 351
by Blasch E., Dezert J., Valin P.

The effective use of the DSmT for multi-class classification / 359
by Abbas N., Chibani Y., Martin A., Smarandache F.

Intelligent Alarm Prioritization based on DSmT / 381
by Tchamova A., Dezert J.

Human Integration of Motion and Texture Information in Visual Slant Estimation / 387
by Tchamova A., Dezert J., Bocheva N., Konstantinova P.

Overview of Dempster-Shafer and Belief Function Tracking Methods / 393
by Blasch E., Dezert J., Pannetier B.

Tracking applications with fuzzy-based fusion rules / 403
by Tchamova A., Dezert J.

URREF reliability versus credibility in information fusion (STANAG 2511) / 409
by Blasch E., Laskey K.B., Jousselme A.-L., Dragos V., Costa P.C.G., Dezert J.

Application of New Absolute and Relative Conditioning Rules in Threat Assessment / 417
by Krenc K., Smarandache F.

A DSmT Based Combination Systems for Handwritten Signature Verification / 423
by Abbas N., Chibani Y., Hadjadji B.

Automatic Aircraft Recognition using DSmT and HMM / 441
by Li X.-D., Pan J.-D., Dezert J.

Static versus Dynamic Data Information Fusion analysis using DDDAS for Cyber Security Trust / 449
by Blasch E., Al-Nashif Y., Hariri S.

Part 3: List of References related with DSmT / 463

Publications / 465
Seminars on DSmT / 478
M.Sc. and Ph.D Theses related with DSmT / 449
Awards / 485
Patent / 485

Biographies of contributors / 487

Foreword

The fourth volume on *Advances and Applications of Dezert-Smarandache Theory* (DSmT) for information fusion collects theoretical and applied contributions of researchers working in different fields of applications and in mathematics. The contributions (see *List of Articles published in this book*, at the end of the volume) have been published or presented after disseminating the third volume (2009, http://fs.gallup.unm.edu/DSmT-book3.pdf) in international conferences, seminars, workshops and journals.

First Part of this book presents the theoretical advancement of DSmT, dealing with Belief functions, conditioning and deconditioning, Analytic Hierarchy Process, Decision Making, Multi-Criteria, evidence theory, combination rule, evidence distance, conflicting belief, sources of evidences with different importance and reliabilities, importance of sources, pignistic probability transformation, Qualitative reasoning under uncertainty, Imprecise belief structures, 2-Tuple linguistic label, Electre Tri Method, hierarchical proportional redistribution, basic belief assignment, subjective probability measure, neutrosophic logic, Evidence theory, outranking methods, Dempster-Shafer Theory, Bayes fusion rule, frequentist probability, mean square error, controlling factor, optimal assignment solution, data association, Transferable Belief Model, and others.

More applications of DSmT have emerged in the past years since the apparition of the third book of DSmT 2009. Subsequently, the second part of this volume is about applications of DSmT in correlation with Electronic Support Measures, belief function, sensor networks, Ground Moving Target and Multiple target tracking, Vehicle-Born Improvised Explosive Device, Belief Interacting Multiple Model filter, seismic and acoustic sensor, Support Vector Machines, Alarm classification, ability of human visual system, Uncertainty Representation and Reasoning Evaluation Framework, Threat Assessment, Handwritten Signature Verification, Automatic Aircraft Recognition, Dynamic Data-Driven Application System, adjustment of secure communication trust analysis, and so on.

Finally, the third part presents a *List of References related with DSmT* published or presented along the years since its inception in 2004, chronologically ordered.

We want to thank all the contributors of this fourth volume for their research works and their interests in the development of DSmT.

We are grateful as well to other colleagues for encouraging us to edit a new volume, for sharing with us several ideas and for their questions and comments on DSmT through the years. We thank the International Society of Information Fusion (www.isif.org) for diffusing main research works related to information fusion (including DSmT) in the international fusion conferences series over the years.

This book is dedicated to the memory of our good friends and colleagues Dr. Jean-Pierre Le Cadre, Prof. Pierre Valin (ISIF president 2006) and Prof. Darko Mušicki (ISIF President 2008) who have always been very active in ISIF and in the organization of past fusion conferences. We will never forget them.

Also, Florentin Smarandache is grateful to The University of New Mexico, U.S.A., that many times partially sponsored him to attend international conferences, workshops and seminars on Information Fusion, and Jean Dezert is grateful to the Department of Information Modeling and Processing (DTIM) at the French Aerospace Lab (Office National d'Etudes et de Recherches Aérospatiales), Palaiseau, France, for encouraging him to carry on this research and for its financial support.

For the next volume, the authors are pleased to send their articles on DSmT to the editors:

Prof. Florentin Smarandache (fsmarandache@gmail.com)
Dr. Jean Dezert (jdezert@gmail.com).

The Editors.

Part 1:

Theoretical advances on DSmT

Non Bayesian Conditioning and Deconditioning

Jean Dezert
Florentin Smarandache

Originally published as: Dezert J., Smarandache F.- *Non Bayesian conditioning and deconditioning*, in Proc. of International Workshop on Belief Functions, Brest, France, April 2-4, 2010, and reprinted with permission.

***Abstract*—In this paper, we present a Non-Bayesian conditioning rule for belief revision. This rule is truly Non-Bayesian in the sense that it doesn't satisfy the common adopted principle that when a prior belief is Bayesian, after conditioning by X, $Bel(X|X)$ must be equal to one. Our new conditioning rule for belief revision is based on the proportional conflict redistribution rule of combination developed in DSmT (Dezert-Smarandache Theory) which abandons Bayes' conditioning principle. Such Non-Bayesian conditioning allows to take into account judiciously the level of conflict between the prior belief available and the conditional evidence. We also introduce the deconditioning problem and show that this problem admits a unique solution in the case of Bayesian prior; a solution which is not possible to obtain when classical Shafer and Bayes conditioning rules are used. Several simple examples are also presented to compare the results between this new Non-Bayesian conditioning and the classical one.**

Keywords: Belief functions, conditioning, deconditioning, probability, DST, DSmT, Bayes rule.

I. Introduction

The question of the updating of probabilities and beliefs has yielded, and still yields, passionate philosophical and mathematical debates [3], [6], [7], [9], [12], [13], [17], [20], [22] in the scientific community and it arises from the different interpretations of probabilities. Such question has been reinforced by the emergence of the possibility and the evidence theories in the eighties [4], [16] for dealing with uncertain information. We cannot browse in details here all the different authors' opinions [1], [2], [8], [10], [14], [15] on this important question but we suggest the reader to start with Dubois & Prade survey [5]. In this paper, we propose a true Non-Bayesian rule of combination which doesn't satisfy the well-adopted Bayes principle stating that $P(X|X) = 1$ (or $Bel(X|X) = 1$ when working with belief functions). We show that by abandoning such Bayes principle, one can take into account more efficiently in the conditioning process the level of the existing conflict between the prior evidence and the new conditional evidence. We show also that the full deconditioning is possible in some specific cases. Our approach is based on belief functions and the Proportional Conflict Redistribution (mainly PCR5) rule of combination developed in Dezert-Smarandache Theory (DSmT) framework [18]. Why we use PCR5 here? Because PCR5 is very efficient to combine conflicting sources of evidences[1] and because Dempster's rule often considered as a generalization of Bayes rule is actually not deconditionable (see examples in the sequel), contrariwise to PCR5, that's why we utilize PCR5. This paper is organized as follows. In section II, we briefly recall Dempster's rule of combination and Shafer's Conditioning Rule (SCR) proposed in Dempster-Shafer Theory (DST) of belief functions [16]. In section III, we introduce a new Non-Bayesian conditioning rule and show its difference with respect to SCR. In section IV, we introduce the dual problem, called the deconditioning problem. Some examples are given in section V with concluding remarks in section VI.

II. Shafer's conditioning rule

In DST, a normalized basic belief assignment (bba) $m(.)$ is defined as a mapping from the power set 2^Θ of the finite discrete frame of discernment Θ into $[0,1]$ such that $m(\emptyset) = 0$ and $\sum_{X\in 2^\Theta} m(X) = 1$. Belief and plausibility functions are in one-to-one correspondence with $m(.)$ and are respectively defined by $Bel(X) = \sum_{Z\in 2^\Theta, Z\subseteq X} m(Z)$ and $Pl(X) = \sum_{Z\in 2^\Theta, Z\cap X\neq 0} m(Z)$. They are usually interpreted as lower and upper bounds of a unknown measure of subjective probability $P(.)$, i.e. $Bel(X) \leq P(X) \leq Pl(X)$ for any X. In DST, the combination of two independent sources of evidence characterized by $m_1(.)$ and $m_2(.)$ is done using Dempster's rule as follows[2]:

$$m_{DS}(X) = \frac{\sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2 = X}} m_1(X_1)m_2(X_2)}{1 - \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2 = \emptyset}} m_1(X_1)m_2(X_2)} \quad (1)$$

Shafer's conditioning rule[3] (SCR) is obtained as the result of Dempster's combination of the given prior bba $m_1(.)$ with the conditional evidence, say Y represented by a source $m_2(.)$ only focused on Y, that is such that $m_2(Y) = 1$. In other words, $m(X|Y) = m_{DS}(X) = (m_1 \oplus m_2)(X)$ using $m_2(Y) = 1$ and where $\oplus$ symbol denotes here Dempster's

[1] Due to space limitation, we do not present, nor justify again PCR5 w.r.t. other rules since this has been widely explained in the literature with many examples and discussions, see for example [18], Vol. 2. and our web page.

[2] assuming that the numerator is not zero (the sources are not in total conflict).

[3] also called Dempster's conditioning by Glenn Shafer in [16].

fusion rule (1). It can be shown [16] that the conditional belief and the plausibility are given by[4]:

$$Bel(X|Y) = \sum_{\substack{Z\in 2^\Theta \\ Z\subseteq X}} m_{DS}(Z|Y) = \frac{Bel_1(X\cup \bar{Y}) - Bel_1(\bar{Y})}{1 - Bel_1(\bar{Y})} \quad (2)$$

$$Pl(X|Y) = \sum_{\substack{Z\in 2^\Theta \\ Z\cap X\neq \emptyset}} m_{DS}(Z|Y) = \frac{Pl_1(X\cap Y)}{Pl_1(Y)} \quad (3)$$

When the belief is Bayesian[5], i.e. $Bel(.|Y) = Pl(.|Y) = P(.|Y)$, SCR reduces to classical conditional probability definition (Bayes formula), that is $P(X|Y) = P(X\cap Y)/P(Y)$, with $P(.) = m_1(.)$. Note that when $Y = X$ and as soon as $Bel(\bar{X}) < 1$, one always gets from (2), $Bel(X|X) = 1$ because $Bel_1(X\cup\bar{Y}) = Bel_1(X\cup\bar{X}) = Bel_1(\Theta) = 1$. For Bayesian belief, this implies $P(X|X) = 1$ for any X such that $P_1(X) > 0$, which we call *Bayes principle*. Other alternatives have been proposed in the literature [8], [15], [21], but almost all of them satisfy Bayes principle and they are all somehow extensions/generalization of Bayes rule. A true Non-Bayesian conditioning (called weak conditioning) was however introduced by Planchet in 1989 in [14] but it didn't bring sufficient interest because Bayes principle is generally considered as the best solution for probability updating based on different arguments for supporting such idea. Such considerations didn't dissuade us to abandon Bayes principle and to explore new Non-Bayesian ways for belief updating, as Planchet did in nineties. We will show in next section why Non-Bayesian conditioning can be interesting.

III. A Non Bayesian Conditioning Rule

Before presenting our Non Bayesian Conditioning Rule, it is important to recall briefly the Proportional Conflict Redistribution Rule no. 5 (PCR5) which has been proposed as a serious alternative of Dempster's rule [16] in Dezert-Smarandache Theory (DSmT) [18] for dealing with conflicting belief functions. In this paper, we assume working in the same fusion space as Glenn Shafer, i.e. on the power set 2^Θ of the finite frame of discernment Θ made of exhaustive and exclusive elements.

A. PCR5 rule of combination

Definition: Let's $m_1(.)$ and $m_2(.)$ be two independent[6] bba's, then the PCR5 rule of combination is defined as follows (see [18], Vol. 2 for details, justification and examples) when working in power set 2^Θ: $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2 = X}} m_1(X_1)m_2(X_2) + \sum_{\substack{X_2\in 2^\Theta \\ X_2\cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X)+m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X)+m_1(X_2)}\right] \quad (4)$$

All fractions in (4) having zero denominators are discarded. The extension and a variant of (4) (called PCR6) for combining $s > 2$ sources and for working in other fusion spaces is presented in details in [18]. Basically, in PCR5 the partial conflicting masses are redistributed proportionally to the masses of the elements which are involved in the partial conflict only, so that the specificity of the information is entirely preserved through this fusion process. It has been clearly shown in [18], Vol. 3, chap. 1 that Smets' rule[7] is not so useful, nor cogent because it doesn't respond to new information in a global or in a sequential fusion process. Indeed, very quickly Smets fusion result commits the full of mass of belief to the empty set!!! In applications, some ad-hoc numerical techniques must be used to circumvent this serious drawback. Such problem doesn't occur with PCR5 rule. By construction, other well-known rules like Dubois & Prade, or Yager's rule, and contrariwise to PCR5, increase the non-specificity of the result.

Properties of PCR5:

- (P0): PCR5 rule is not associative, but it is quasi-associative (see [18], Vol. 2).
- (P1): PCR5 Fusion of two non Bayesian bba's is a non Bayesian bba.
  Example: Consider $\Theta = \{A, B, C\}$ with Shafer's model and with the two non Bayesian bba's $m_1(.)$ and $m_2(.)$ given in Table I. The PCR5 fusion result (rounded at the fourth decimal) is given in the right column of the Table I. One sees that $m_{PCR5}(.)$ in a non Bayesian bba since some of its focal elements are not singletons.

Table I
PCR5 fusion of two non Bayesian bba's.

Focal Elem.	$m_1(.)$	$m_2(.)$	$m_{PCR5}(.)$
A	0.1	0.2	0.3850
B	0.2	0.1	0.1586
C	0.1	0.2	0.1990
$A\cup B$	0.3	0	0.0360
$A\cup C$	0	0.5	0.2214
$A\cup B\cup C$	0.3	0	0

- (P2): PCR5 Fusion of a Bayesian bba with a non Bayesian bba is a non Bayesian bba in general[8].
  Example: Consider $\Theta = \{A, B, C\}$ with Shafer's model and Bayesian and a non Bayesian bba's $m_1(.)$ and $m_2(.)$ to combine as given in Table II. The PCR5 fusion result is given in the right column of the Table II. One sees that $m_{PCR5}(.)$ is a non Bayesian bba since some of its focal elements are not singletons.
  This property is in opposition with Dempster's rule property (see Theorem 3.7 p. 67 in [16]) which states that if Bel_1 is Bayesian and if Bel_1 and Bel_2 are combinable, then Dempster's rule provides always a Bayesian belief function. The result of Dempster's rule noted $m_{DS}(.)$ for

[4] $\bar{Y}$ denotes the complement of Y in the frame Θ.
[5] the focal elements of $m_1(.|Y)$ are singletons only.
[6] i.e. each source provides its bba independently of the other sources.
[7] i.e. the non normalized Dempster's rule.
[8] In some cases, it happens that Bayesian $\oplus$ Non-Bayesian $=$ Bayesian. For example, with $\Theta = \{A, B, C\}$, Shafer's model, $m_1(A) = 0.3$, $m_1(B) = 0.7$ and $m_2(A) = 0.1$, $m_2(B) = 0.2$, $m_2(C) = 0.4$ and $m_2(A\cup B) = 0.3$, one gets $m_{PCR5}(A) = 0.2162$, $m_{PCR5}(B) = 0.6134$ and $m_{PCR5}(C) = 0.1704$ which is a Bayesian bba.

Table II
PCR5 FUSION OF BAYESIAN AND NON BAYESIAN BBA'S.

Focal Elem.	$m_1(.)$	$m_2(.)$	$m_{DS}(.)$	$m_{PCR5}(.)$
A	0.1	0	0.0833	0.0642
B	0.2	0.3	0.1000	0.1941
C	0.7	0.2	0.8167	0.6703
$A \cup C$	0	0.5	0	0.0714

this example is given in Table II for convenience. This is the major difference between PCR5 and Dempster's rule, not to mention the management of conflicting information in the fusion process of course.
In summary, and using $\oplus$ symbol to denote the generic fusion process, one has

– **With Dempster's rule** :

$$\text{Bayesian} \oplus \text{Non-Bayesian} = \text{Bayesian}$$

– **With PCR5 rule**:

$$\text{Bayesian} \oplus \text{Non-Bayesian} = \text{Non-Bayesian (in general)}$$

- (P3): PCR5 Fusion of two Bayesian bba's is a Bayesian bba (see [18], Vol. 2, pp. 43–45 for proof).

  Example: $\Theta = \{A, B, C\}$ with Shafer's model and let's consider Bayesian bba's given in the next Table. The result of PCR5 fusion rule is given in the right column of Table III. One sees that $m_{PCR5}(.)$ is Bayesian since its focal elements are singletons of the fusion space 2^Θ.

Table III
PCR5 FUSION OF TWO BAYESIAN BBA'S.

Focal Elem.	$m_1(.)$	$m_2(.)$	$m_{DS}(.)$	$m_{PCR5}(.)$
A	0.1	0.4	0.0870	0.2037
B	0.2	0	0	0.0567
C	0.7	0.6	0.9130	0.7396

B. A true Non Bayesian conditioning rule

Here[9] we follow the footprints of Glenn Shafer in the sense that we consider the conditioning as the result of the fusion of any prior mass $m_1(.)$ defined on 2^Θ with the bba $m_2(.)$ focused on the conditional event $Y \neq \emptyset$, i.e. $m_2(Y) = 1$. We however replace Dempster's rule by the more efficient[10] Proportional Conflict Redistribution rule # 5 (PCR5) given by (4) proposed in DSmT [18]. This new conditioning rule is not Bayesian and we use the symbol $\parallel$ (parallel) instead of classical symbol $|$ to avoid confusion in notations. Let's give the expression of $m(X \parallel Y)$ resulting of the PCR5 fusion of any prior bba $m_1(.)$ with $m_2(.)$ focused on Y. Applying (4):

$$m(X \parallel Y) = S_1^{\text{PCR5}}(X,Y) + S_2^{\text{PCR5}}(X,Y) + S_3^{\text{PCR5}}(X,Y) \quad (5)$$

with

$$S_1^{\text{PCR5}}(X,Y) \triangleq \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) \quad (6)$$

$$S_2^{\text{PCR5}}(X,Y) \triangleq m_1(X)^2 \sum_{\substack{X_2 \in 2^\Theta \\ X \cap X_2 = \emptyset}} \frac{m_2(X_2)}{m_1(X) + m_2(X_2)} \quad (7)$$

$$S_3^{\text{PCR5}}(X,Y) \triangleq m_2(X)^2 \sum_{\substack{X_2 \in 2^\Theta \\ X \cap X_2 = \emptyset}} \frac{m_1(X_2)}{m_2(X) + m_1(X_2)} \quad (8)$$

where $m_2(Y) = 1$ for a given $Y \neq \emptyset$.

Since Y is the single focal element of $m_2(.)$, the term $S_1^{\text{PCR5}}(X,Y)$ in (5) is given by $\sum_{\substack{X_1 \in 2^\Theta \\ X_1 \cap Y = X}} m_1(X_1)$, the term $S_2^{\text{PCR5}}(X,Y)$ equals $\delta(X \cap Y = \emptyset) \cdot \frac{m_1(X)^2}{1+m_1(X)}$, and the term $S_3^{\text{PCR5}}(X,Y)$ can be expressed depending on the value of X with respect to the conditioning term Y:

- If $X \neq Y$ then $m_2(X \neq Y) = 0$ (by definition), and thus $S_3^{\text{PCR5}}(X,Y) = 0$.
- If $X = Y$ then $m_2(X = Y) = 1$ (by definition), and thus $S_3^{\text{PCR5}}(X,Y) = \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap Y = \emptyset}} \frac{m_1(X_2)}{1+m_1(X_2)}$

Finally, $S_3^{\text{PCR5}}(X,Y)$ can be written as

$$\begin{aligned} S_3^{\text{PCR5}}(X,Y) &= \underbrace{\delta(X \neq Y) \cdot 0}_{0} + \delta(X = Y) \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap Y = \emptyset}} \frac{m_1(X_2)}{1 + m_1(X_2)} \\ &= \delta(X = Y) \cdot \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap Y = \emptyset}} \frac{m_1(X_2)}{1 + m_1(X_2)} \end{aligned}$$

Finally, $m(X \parallel Y)$ for $X \neq \emptyset$ and $Y \neq \emptyset$ are given by

$$\begin{aligned} m(X \parallel Y) = &\sum_{\substack{X_1 \in 2^\Theta \\ X_1 \cap Y = X}} m_1(X_1) + \delta(X \cap Y = \emptyset) \cdot \frac{m_1(X)^2}{1 + m_1(X)} \\ &+ \delta(X = Y) \cdot \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap Y = \emptyset}} \frac{m_1(X_2)}{1 + m_1(X_2)} \quad (9) \end{aligned}$$

$m(\emptyset \parallel Y \neq \emptyset) = 0$ by definition, since PCR5 fusion doesn't commit mass on the empty set. $m(X \parallel \emptyset)$ is kept undefined[11] since it doesn't make sense to revise a bba by an impossible event. Based on the classical definitions of $Bel(.)$ and $Pl(.)$ functions [16], one has:

$$Bel(X \parallel Y) = \sum_{\substack{Z \in 2^\Theta \\ Z \subseteq X}} m(Z \parallel Y) \quad (10)$$

$$Pl(X \parallel Y) = \sum_{\substack{Z \in 2^\Theta \\ Z \cap X \neq \emptyset}} m(Z \parallel Y) \quad (11)$$

The "true" unknown (non Bayesian) conditional subjective probability, denoted $P(X||Y)$, must satisfy

$$Bel(X \parallel Y) \leq P(X||Y) \leq Pl(X||Y) \quad (12)$$

[9]More sophisticated conditioning rules have been proposed in [18], Vol. 2.

[10]It deals better with partial conflicts than other rules unlike Dempster's rule, it does not increase the non-specificity of the result unlike Dubois & Prade or Yager's rule, and it does respond to new information unlike Smets rule.

[11]One could also define $m(\emptyset \parallel \emptyset) = 1$ and $m(X \neq \emptyset \parallel \emptyset) = 0$ which however would not be a normal bba.

$P(X||Y)$ can be seen as an imprecise probability and used within IPT (Imprecise Probability Theory) [23] if necessary, or can be approximated from $m(.||Y)$ using some probabilistic transforms, typically the pignistic transform [19] or the DSmP transform [18] (Vol.3, Chap. 3). The search for direct close-form expressions of $Bel(X \parallel Y)$ and $Pl(X \parallel Y)$ from $Bel_1(.)$ and $Pl_1(.)$ appears to be an open difficult problem.

IV. Deconditioning

In the previous section we have proposed a new non Bayesian conditioning rule based on PCR5. This rule follows Shafer's idea except that we use PCR5 instead of Dempster's rule because we have shown the better efficiency of PCR5 to deal with conflicting information w.r.t. other rules. In this section, we also show the great benefit of such PCR5 rule for the deconditioning problem. The belief conditioning problem consists in finding a way to update any prior belief function ($Bel(.)$, $Pl()$ or $m(.)$) with a new information related with the (belief of) occurrence in a given conditional proposition of the fusion space, say Y, in order to get a new belief function called conditional belief function. The deconditioning problem is the inverse (dual) problem of conditioning. It consists to retrieve the prior belief function from a given posterior/conditional belief function. Deconditioning has not been investigated in deep so far in the literature (to the knowledge of the authors) since is is usually considered as impossible to achieve[12], it may present great interest for applications in advanced information systems when only a posterior belief is available (say provided by an human or an AI-expert system), but for some reason we need to compute a new conditioning belief based on a different conditional hypothesis. This motivates our research for developing deconditioning techniques. Since $Bel(.)$, $Pl()$ are in one-to-one correspondence with the basic belief assignment (bba) mass $m(.)$, we focus our analysis on the deconditioning of the conditional bba. More simply stated, we want to see if for any given conditional bba $m(.||Y)$ we can compute $m_1(.)$ such that $m(.||Y) = PCR5(m_1(.), m_2(.))$ with $m_2(Y) = 1$ and where $PCR5(m_1(.), m_2(.))$ denotes the PCR5 fusion of $m_1(.)$ with $m_2(.)$. Let's examine the two distinct cases for the deconditiong problem depending on the (Bayesian or non-Bayesian) nature of the prior $m_1(.)$.

- **Case of Bayesian prior** $m_1(.)$: Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, with $n \geq 2$, Shafer's model, where all θ_i are singletons. Let $m_1 : \Theta \mapsto [0,1]$ be a Bayesian bba/mass. In that case, the deconditioning problem admits a unique solution and we can always compute $m_1(.)$ from $m(.||Y)$ but two distinct cases must be analyzed depending on the cardinality of the conditional term Y.
  Case 1: When Y is a singleton, i.e. $|Y| = 1$. Suppose $m_2(Y) = 1$, with $Y = \theta_{j_0}$, for $j_0 \in \{1, 2, \ldots, n\}$, where j_0 is fixed. Since the bba's $m_1(.)$ and $m_2(.)$ are both Bayesian in this case, $m(.||Y)$ is also a Bayesian bba (property P3), therefore $m(\theta_i||Y) = a_i$, where all $a_i \in [0,1]$ with $\sum_{i=1}^{n} a_i = 1$. How to find $m_1(.)$ such that $m(.||Y) = PCR5(m_1(.), m_2(.))$? Let's denote $m_1(\theta_i) = x_i$, where all $x_i \in [0,1]$ and $\sum_{i=1}^{n} x_i = 1$. We need to find all these x_i. We now combine $m_1(.)$ with $m_2(.)$ using PCR5 fusion rule. We transfer x_i, for $\forall i \neq j_0$, to θ_i and θ_{j_0} proportionally with respect to their corresponding masses, x_i and 1 respectively: $\frac{w_{\theta_i}}{x_i} = \frac{w_{\theta_{j_0}}}{1} = \frac{x_i}{x_i+1}$ whence $w_{\theta_i} = \frac{x_i^2}{x_i+1}$ and $w_{\theta_{j_0}} = \frac{x_i}{x_i+1}$, while $\alpha_{j_0} = x_{j_0} + \sum_{\substack{i=1 \\ i \neq j_0}}^{n} \frac{x_i}{x_i+1}$ or $\alpha_{j_0} = 1 - \sum_{\substack{i=1 \\ i \neq j_0}}^{n} \frac{x_i^2}{x_i+1}$. Since we need to find all unknowns x_i, $i = 1, \ldots, n$, we need to solve $\frac{x_i^2}{x_i+1} = a_i$, for $i \neq j_0$ for x_i; since $\alpha_{j_0} = x_{j_0} + \sum_{\substack{i=1 \\ i \neq j_0}}^{n} \frac{x_i}{x_i+1} = a_{j_0}$, we get $x_{j_0} = a_{j_0} - \sum_{\substack{i=1 \\ i \neq j_0}}^{n} \frac{x_i}{x_i+1} = 1 - \sum_{\substack{i=1 \\ i \neq j_0}}^{n} x_i$.
  Case 2: When Y is not a singleton, i.e. $|Y| > 1$ (Y can be a partial or total ignorance). Suppose $m_2(Y) = 1$, with $Y = \theta_{j_1} \cup \theta_{j_2} \cup \ldots \cup \theta_{j_p}$, where all j_1, j_2, ..., j_p are different and they belong to $\{1, 2, \ldots, n\}$, $2 \leq p \leq n$. We keep the same notations for $m(.||Y)$ and Bayesian $m_1(.)$. The set $\{j_1, j_2, \ldots, j_p\}$ is denoted J for notation convenience. Similarly, using PCR5 rule we transfer x_i, $\forall i \notin J$, to x_i and to the ignorance $Y = \theta_{j_1} \cup \ldots \cup \theta_{j_p}$ proportionally with respect to x_i and 1 respectively (as done in case 1). So, x_i for $i \notin J$ is found from solving the equation $\frac{x_i^2}{x_i+1} = a_i$, which gives[13] $x_i = (a_i + \sqrt{a_i^2 + 4a_i})/2$; and $x_{j_r} = a_{j_r}$ for $r \in \{1, 2, \ldots, p\}$.
- **Case of Non-Bayesian prior** $m_1(.)$:
  Unfortunately, when $m_1(.)$ is Non-Bayesian, the (PCR5-based) deconditioning problem doesn't admit one unique solution in general (see the example 2.1 in the next section). But the method used to decondition PCR5 when $m_1(.)$ is Bayesian can be generalized for $m_1(.)$ non-Bayesian in the following way: 1) We need to know the focal elements of $m_1(.)$, then we denote the masses of these elements by say x_1, x_2, ..., x_n; 2)Then we combine using the conjunctive rule $m_1(.)$ with $m_2(Y) = 1$, where Y can be a singleton or an ignorance; 3) Afterwards, we use PCR5 rule and we get some results like: $f_i(x_1, ..., x_n)$ for each element, where $i = 1, 2, \ldots$. Since we know the results of PCR5 as $m(.||Y) = a_i$ for each focal element, then we form a system of non-linear equations: $f_i(x_1, x_2, ..., x_n) = a_i$ and we need to solve it. Such systems of equations however can admit several solutions. We can select a solution satisfying an additional criterion like by example the minimum (or the maximum) of specificity depending of the kind of Non-Bayesian prior we need to use.

V. Examples

A. Example 1: Conditioning of a Bayesian prior belief

Let's consider $\Theta = \{A, B, C\}$, Shafer's model, and the prior bba's $m_1(.)$ and $m_1'(.)$ given in Table IV and the conditional evidence $Y = A \cup B$.

[12] This truly happens when classical Bayes conditioning is used.

[13] The solution $x_i = (a_i - \sqrt{a_i^2 + 4a_i})/2$ must be discarded since it is negative and cannot be considered as a mass of belief.

Table IV
BAYESIAN PRIORS (INPUTS).

Focal Elem.	m_1	$m'_1(.)$
A	0.49	0.01
B	0.49	0.01
C	0.02	0.98

The significance of having two cases in the Bayesian prior case is straighforward. We just want to show that two different priors can yield to the same posterior bba with Bayes/SCR rule and thus we cannot retrieve these two distinct priors cases from the posterior bba. We show that the total deconditioning is possible however when using our non-Bayesian conditioning rule. SCR and PCR5-based conditioning of $m_1(.)$ and $m'_1(.)$ are given[14] in Table V. One sees that SCR of the two distinct bba's $m_1(.)$ and $m'_1(.)$ yield the same posterior/conditional bba $m(.|Y)$ which means that in this very simple Bayesian prior case, the deconditioning of $m(.|Y)$ is impossible to obtain since at least two solutions[15] for the prior beliefs are admissible. The results provided by PCR5-based conditioning makes more sense in authors' point of view since it better takes into account the degree of conflicting information in the conditioning process. One sees that two distinct Bayesian priors yield two distinct posterior bba's with PCR5-based conditioning. If one examines the belief and plausibility functions, one gets, using notation $\Delta(.|Y) = [Bel(.|Y), Pl(.|Y)]$, $\Delta'(.|Y) = [Bel'(.|Y), Pl'(.|Y)]$, $\Delta(.||Y) = [Bel(.||Y), Pl(.||Y)]$ and $\Delta'(.||Y) = [Bel'(.||Y), Pl'(.||Y)]$:

Table V
CONDITIONAL BBA'S.

Focal Elem.	$m(.\|Y)$	$m'(.\|Y)$	$m(.\|\|Y)$	$m'(.\|\|Y)$
A	0.5	0.5	0.4900	0.0100
B	0.5	0.5	0.4900	0.0100
C	0	0	0.00039215	0.48505051
$A \cup B$	0	0	0.01960785	0.49494949

Table VI
CONDITIONAL LOWER AND UPPER BOUNDS OF CONDITIONAL PROBABILITIES

2^Θ	$\Delta(.\|Y) = \Delta'(.\|Y)$	$\Delta(.\|\|Y)$	$\Delta'(.\|\|Y)$
$\emptyset$	[0,0]	[0,0]	[0,0]
A	[0.5,0.5]	[0.4900, 0.5096]	[0.0100, 0.5050]
B	[0.5,0.5]	[0.4900, 0.5096]	[0.0100, 0.5050]
C	[0,0]	[0.0004, 0.0004]	[0.4850,0.4850]
$Y = A \cup B$	[1,1]	[0.9996,0.9996]	[0.5150,0.5150]
$A \cup C$	[0.5,0.5]	[0.4904, 0.5100]	[0.4950, 0.9900]
$B \cup C$	[0.5,0.5]	[0.4904, 0.5100]	[0.4950, 0.9900]
$A \cup B \cup C$	[1,1]	[1,1]	[1,1]

The interval $\Delta(.|Y)$ corresponds to lower and upper bounds of conditional subjective probabilities $P(.|Y)$ and $\Delta(.||Y)$ corresponds to lower and upper bounds of $P(.||Y)$ (similarly for $\Delta'(.|Y)$ and $\Delta'(.||Y)$). From the Table VI, one sees that the property P2 is verified and we get an imprecise conditional probability. One sees that contrariwise to SCR (equivalent to Bayes rule in this case), one gets $Bel(Y||Y) < 1$ and also $Pl(Y||Y) < 1$. $\Delta(.||Y)$ and $\Delta'(.||Y)$ are very different because priors were also very different. This is an appealing property. If one approximates[16] the conditional probability by the mid-value of their lower and upper bounds[17], one gets values given in Table VII.

Table VII
CONDITIONAL APPROXIMATE SUBJECTIVE PROBABILITIES.

2^Θ	$P(.\|Y) = P'(.\|Y)$	$P(.\|\|Y)$	$P'(.\|\|Y)$
$\emptyset$	0	0	0
A	0.5	0.4998	0.2575
B	0.5	0.4998	0.2575
C	0	0.0004	0.4850
$Y = A \cup B$	1	0.9996	0.5150
$A \cup C$	0.5	0.5002	0.7425
$B \cup C$	0.5	0.5002	0.7425
$A \cup B \cup C$	1	1	1

When the conditioning hypothesis supports the prior belief (as for $m_1(.)$ and $m_2(.)$ which are in low conflict) the PCR5-based conditioning reacts as SCR (as Bayes rule when dealing with Bayesian priors) and $P(X.||Y)$ is very close to $P(.|Y)$. When the prior and the conditional evidences are highly conflicting (i.e. like $m'_1(.)$ and $m_2(.)$, PCR5-based conditioning rule is much more prudent than Shafer's rule and that's why it allows the possibility to have $P(Y||Y) < 1$. Such property doesn't violate the fundamental axioms (nonnegativity, unity and additivity) of Kolmogorov axiomatic theory of probabilities and this can be verified easily in our example. In applications, it is much better to preserve all available information and to work directly with conditional bba's whenever possible rather than with approximate subjective conditional probabilities.

The deconditioning of the posterior bba's $m(. \parallel Y)$ given in the Table V is done using the principle described in section IV (when $m_1(.)$ is assumed Bayesian and for case 2). We denote the unknowns $m_1(A) = x_1$, $m_1(B) = x_2$ and $m_1(C) = x_3$. Since $Y = A \cup B$ and $J = \{1, 2\}$, we solve the following system of equations (with the constraint $x_i \in [0, 1]$): $x_1 = a_1 = 0.49$, $x_2 = a_2 = 0.49$ and $x_3^2/(x_3 + 1) = a_3 = 0.00039215$. Therefore, one gets after deconditioning $m_1(A) = 0.49$, $m_1(B) = 0.49$ and $m_1(C) = 0.02$. Similarly, the deconditioning of $m'(. \parallel Y)$ given in the Table V yields $m'_1(A) = 0.01$, $m'_1(B) = 0.01$ and $m'_1(C) = 0.98$.

Note that, contrarywise to Bayes or to Jeffrey's rules [8], [11], [21], it is possible to update the prior opinion about an event A even if $P(A) = 0$ using this Non-Bayesian rule. For example, let's consider $\Theta = \{A, B, C\}$, Shafer's model and the prior Bayesian mass $m_1(A) = 0$, $m_1(B) = 0.3$ and $m_1(C) = 0.7$, i.e. $Bel_1(A) = P_1(A) = Pl(A) = 0$. Assume that the conditional evidence is $Y = A \cup B$, then one gets with SCR $m(B|A \cup B) = 1$ and with PCR5-based conditioning $m(B \parallel A \cup B) = 0.30$, $m(A \cup B \parallel A \cup B) = 0.41176$ and $m(C \parallel A \cup B) = 0.28824$, which means that $P(A|A \cup B) = 0$ with SCR/Bayes rule (i.e. no update on A), whereas $[Bel(A \parallel A \cup B), Pl(A \parallel A \cup B)] = [0, 0.41176]$, $[Bel(B \parallel$

[14] Due to space limitation constraints, the verification is left to the reader.
[15] Actually an infinite number of solutions exists.
[16] When the lower bound is equal to the upper bound, one gets the exact probability value.
[17] More sophisticated transformations could be used instead as explained in [18], Vol. 3.

$A \cup B), Pl(B \parallel A \cup B)] = [0.30, 0.71176]$ and $[Bel(C \parallel A \cup B), Pl(C \parallel A \cup B)] = [0.28823, 0.28823]$, that is $P(A \parallel A \cup B) \in [0, 0.41176]$. Typically, if one approximates $P(. \parallel A \cup B)$ by the mid-value of its lower and upper bounds, one will obtain $P(A \parallel A \cup B) = 0.20588$ (i.e. a true update of the prior probability of A), $P(B \parallel A \cup B) = 0.50588$ and $P(C \parallel A \cup B) = 0.28824$.

B. Example 2: Conditioning of a Non-Bayesian prior belief

Example 2.1: Let's consider now $\Theta = \{A, B, C\}$, Shafer's model, the conditioning hypothesis $Y = A \cup B$ and the following Non-Bayesian priors:

Table VIII
NON-BAYESIAN PRIORS (INPUTS).

Focal Elem.	m_1	$m_1'(.)$
A	0.20	0.20
B	0.30	0.30
C	0.10	0.10
$A \cup B$	0.25	0.15
$A \cup B \cup C$	0.15	0.25

The conflict between $m_1(.)$ and $m_2(Y) = 1$ and between $m_1'(.)$ and $m_2(Y) = 1$ is 0.10 in both cases. The results of the conditioning are given in Table IX. One sees that when distinct priors are Non-Bayesian, it can happen that PCR5-based conditioning rule yields also the same posterior bba's. This result shows that in general with Non-Bayesian priors the PCR5-based deconditioning cannot provide a unique solution, unless extra information and/constraints on the prior belief are specified as shown in the next example.

Table IX
CONDITIONAL BBA'S.

Focal Elem.	$m(.\|Y)$	$m'(.\|Y)$	$m(. \parallel Y)$	$m'(. \parallel Y)$
A	0.222	0.222	0.20	0.20
B	0.333	0.333	0.30	0.30
C	0	0	0.01	0.01
$A \cup B$	0.445	0.445	0.49	0.49

Example 2.2: Let's consider now $\Theta = \{A, B, C, D\}$, Shafer's model, the conditional evidence $Y = C \cup D$ and the posterior bba $m(. \parallel C \cup D)$ given in the right column of the table below:

Table X
CONDITIONAL BBA'S.

Focal Elem.	$m_1(.)$	$m_{PCR5}(.)$	$m(. \parallel A)$
A	x_1	$\frac{x_1^2}{1+x_1}$	0.0333
B	x_2	$\frac{x_2^2}{1+x_2}$	0.1667
$C \cup D$	x_3	$x_3 + \frac{x_1}{1+x_1} + \frac{x_2}{1+x_2}$	0.8000

If we assume that the focal elements of the prior bba $m_1(.)$ are the same as for the posterior bba $m(. \parallel C \cup D)$, then with such extra assumption, the deconditioning problem admits a unique solution which is obtained by solving the system of three equations according to Table X; that is $\frac{x_1^2}{1+x_1} = 0.0333$, whence $x_1 \approx 0.2$; $\frac{x_2^2}{1+x_2} = 0.1667$, whence $x_2 \approx 0.5$; $x_3 + \frac{x_1}{1+x_1} + \frac{x_2}{1+x_2} = 0.8000$; whence $x_1 \approx 0.3$. Therefore, the deconditioning of $m(. \parallel C \cup D)$ provides the unique Non-Bayesian solution $m_1(A) = 0.2$, $m_1(B) = 0.5$ and $m_1(C \cup D) = 0.3$.

VI. CONCLUSIONS

In this paper, we have proposed a new Non-Bayesian conditioning rule (denoted $\parallel$) based on the Proportional Conflict Redistribution (PCR) rule of combination developed in DSmT framework. This new conditioning rule offers the advantage to take fully into account the level of conflict between the prior and the conditional evidences for updating belief functions. It is truly Non-Bayesian since it doesn't satisfy Bayes principle because it allows $P(X \parallel X)$ or $Bel(X \parallel X)$ to be less than one. We have also shown that this approach allows to solve the deconditioning (dual) problem for the class of Bayesian priors. More investigations on the deconditioning problem of Non-Bayesian priors need to be done and comparisons of this new rule with respect to the main alternatives of Bayes rule proposed in the literature (typically Jeffrey's rule and its extensions, Planchet's rule, etc) will be presented in details in a forthcoming publication.

REFERENCES

[1] P. Diaconis, S.L. Zabell, *Updating subjective probability*, JASA, 77 380:822-830, 1982

[2] P. Diaconis, S.L. Zabell, *Some alternatives to Bayes's rule*, In Grofman, B., & Owen, G. (Eds.), Proc. Second University of California, Irvine, Conference on Political Economy, pp. 25–38, 1986.

[3] F. Döring, *Why Bayesian psychology is incomplete*, Philosophy of Science 66 (3):389, 1999.

[4] D. Dubois, H. Prade, *Possibility theory: an approach to computerized processing of uncertainty*, Plenum Press, New York, 1988.

[5] D. Dubois, H. Prade, *A survey of belief revision and updating rules in various uncertainty models*, Int. J. Intell. Syst. 9, pp. 61–100, 1994.

[6] D. Eddington, *On conditionals*, Mind, Vol. 104, 404, pp. 235–329, 1995.

[7] P. Grünwald, J. Halpern, *Updating probabilities*, Journal of Artificial Intelligence Research (JAIR) 19, pages 243-278, 2003.

[8] R. Jeffrey, *The logic of decision*, McGraw-Hill, 2nd Edition, 1983.

[9] H.E. Kyburg, Jr., *Bayesian and non-Bayesian evidential updating*, Artif. Intell. 3, pp. 27–294, 1987.

[10] R.H.. Loschi, *Conditioning on uncertain event: extension to Bayesian inference*, Sociedad de Estatistica e Investigacion Operativa, Test, Vol. 11, No. 2, pp. 365-383, 2002.

[11] R.H.. Loschi, P.L. Iglesias, *Jeffrey's rule: an alternative procedure to model uncertainty*, Estadistica 2005, Vol. 57 (168,169), pp. 11–26, 2005.

[12] E. J. Lowe, *Conditional probability and conditional beliefs*, Mind 105 (420), 1996.

[13] E. J. Lowe, *What is 'conditional probability'?*, Analysis 68 (299), pp. 218–223, 2008.

[14] B. Planchet, *Credibility and conditioning*, Journal of Theoretical Probability, 2, pp. 289–299, 1989.

[15] S. Moral, L.M. De Campos, *Updating uncertain information*, Lecture Notes in Computer Science, Springer, Vol. 521, pp. 58–67, 1991.

[16] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[17] G. Shafer, *Conditional probability*, International Statistical Review, 53(3), 261-277, 1985

[18] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vol. 1-3, American Research Press, 2004–2009. http://www.gallup.unm.edu/~smarandache/DSmT.htm

[19] Ph. Smets, *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in AI, Vol. 5, pp. 29–39, 1990.

[20] Ph. Smets, *About updating*, in D'ambrosio B., Smets Ph., and Bonissone P.P. Editors, Uncertainty in AI 91, Morgan Kaufmann, pp. 378–385, 1991.

[21] Ph. Smets, *Jeffrey's rule of conditioning generalized to belief functions*, online paper on Ph. Smets web homepage, July 27th, 1999.

[22] J.H. Sobel, *On the significance of conditional probabilities*, Synthese, Vol. 109 (3), pp. 311–344, 1996.

[23] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, London, 1991.

Multi-criteria decision making based on DSmT-AHP

Jean Dezert
Jean-Marc Tacnet
Mireille Batton-Hubert
Florentin Smarandache

Originally published as: Dezert J, Tacnet J.-M., Batton-Hubert M., Smarandache F. - *Multi-criteria decision making based on DSmT/AHP*, in Proc. of International Workshop on Belief Functions, Brest, France, April 2-4, 2010, and reprinted with permission.

***Abstract*—In this paper, we present an extension of the multi-criteria decision making based on the Analytic Hierarchy Process (AHP) which incorporates uncertain knowledge matrices for generating basic belief assignments (bba's). The combination of priority vectors corresponding to bba's related to each (sub)-criterion is performed using the Proportional Conflict Redistribution rule no. 5 proposed in Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning. The method presented here, called DSmT-AHP, is illustrated on very simple examples.**

Keywords: Analytic Hierarchy Process, AHP, DSmT, Information Fusion, Decision Making, Multi-Criteria.

I. Introduction

The Multi-criteria decision-making (MCDM) problem concerns the elucidation of the level of preferences of decision alternatives through judgments made over a number of criteria [6]. At the Decision-maker (DM) level, a useful method for solving MCDM problem must take into account opinions made under uncertainty and based on distinct criteria with different importances. The difficulty of the problem increases if we consider a group decision-making (GDM) problem involving a panel of decision-makers. Several attempts have been proposed in the literature to solve the MCGDM problem. Among the interesting solutions developed, one must cite the works made by Beynon [3]–[6]. This author developed a method called DS/AHP which extended the Analytic Hierarchy Process (AHP) method of Saaty [15]–[17] with Dempster-Shafer Theory (DST) [23] of belief functions to take into account uncertainty and to manage the conflicts between experts opinions within a hierarchical model approach. In this paper, we propose to follow Beynon's approach, but instead of using DST, we investigate the possibility to use Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning developed since 2002 for overcoming DST limitations[1] [24]. This new approach will be referred as DSmT-AHP method in the sequel. DSmT allows to manage efficiently the fusion of quantitative (or qualitative) uncertain and possibly highly conflicting sources of evidences and proposes new methods for belief conditioning and deconditioning as well [7]. DSmT has been successfully applied in several fields of applications (in defense, medicine, satellite surveillance, biometrics, image processing, etc). In section II, we briefly introduce the principle of the AHP developed by Saaty. In section III, we recall the basis of DSmT and its main rule of combination, called PCR5 (Proportional Conflict Redistribution rule # 5). In section IV, we present the DSmT-AHP method for solving the MCDM problem. The extension of DSmT-AHP method for solving MCGDM problem is then introduced in section V. Conclusions are given in Section VI.

II. The Analytic Hierarchy Process (AHP)

The Analytic Hierarchy Process (AHP) is a structured technique developed by Saaty in [8], [15], [16] based on mathematics and psychology for dealing with complex decisions. AHP and its refinements are used around the world in many decision situations (government, industry, education, healthcare, etc.). It helps the DM to find the decision that best suits his/her needs and his/her understanding of the problem.

[1] A presentation of these limitations with a discussion is done in Chap 1 of [24], Vol. 3. It is shown clearly that the logical refinement proposed by some authors doesn't bring new insights with respect to what is done when working directly on the super-power set (i.e. on the minimal refined frame satisfying Shafer's model). There is no necessity to work with a refined frame in DSmT framework which is very attractive in some real-life problems where the elements of the refined frame do not have any (physical) sense/meaning or are just impossible to clearly determine physically (as a simple example, if Mary and Paul have possibly committed a crime alone or together, there is no way to refine these two persons into three finer exclusive physical elements satisfying Shafer's model). Aside the possibility to deal with different underlying models of the frame, it is worth to note that PCR5 or PCR6 rules provide a better ability than the other rules to deal efficiently with highly conflicting sources of evidences as shown in all fields of applications where they have been tested so far.

AHP provides a comprehensive and rational framework for structuring a decision problem, for representing and quantifying its elements, for relating those elements to overall goals, and for evaluating alternative solutions. The basic idea of AHP is to decompose the decision problem into a hierarchy of more easily comprehended sub-problems, each of which can be analyzed independently. Once the hierarchy is built, the DM evaluates the various elements of the hierarchy by comparing them to one another two at a time [21]. In making the comparisons, the DM can use both objective information about the elements as well as subjective opinions about the elements' relative meaning and importance. The AHP converts these evaluations to numerical values that are processed and compared over the entire range of the problem. A numerical weight or priority is derived for each element of the hierarchy, allowing diverse and often incommensurable elements to be compared to one another in a rational and consistent way. This is the main advantage of AHP with respect to other decision making techniques. At its final step, numerical priorities are calculated for each of the decision alternatives. These numbers represent the alternatives' relative ability to achieve the decision goal. The AHP method can be summarized as [19]:
1) Model the problem as a hierarchy containing the decision goal, the alternatives for reaching it, and the criteria for evaluating the alternatives.
2) Establish priorities among the elements of the hierarchy by making a series of judgments based on pairwise comparisons of the elements.
3) Check the consistency of the judgments and eventually revise the comparison matrices by reasking the experts when the consistency in judgments is too low.
4) Synthesize these judgments to yield a set of overall priorities for the hierarchy.
5) Come to a final decision based on the results of this process.

Example 1: According to his/her own preferences and using the Saaty's 1-9 ordinal scale, a DM wants to buy a car among four available models belonging to the set $\Theta = \{A, B, C, D\}$. To simplify the example, we assume that the objective of DM is to select one of these cars based only on three criteria (C1=Fuel economy, C2=Reliability and C3=Style). According to his/her own preferences, the DM ranks the different criteria pairwise as follows: 1 - Reliability is 3 times as important as fuel economy, 2 - Fuel economy is 4 times as important as style, 3 - Reliability is 5 times as important as style, which means that the DM thinks that Reliability criteria (C2) is the most important criteria, followed by fuel economy (C1) and style is the least important criteria[2]. The relative importance of one criterion over another can be expressed using pairwise comparison matrix (also called knowledge matrix) as follows:

$$\mathbf{M} = \begin{bmatrix} 1/1 & 1/3 & 4/1 \\ 3/1 & 1/1 & 5/1 \\ 1/4 & 1/5 & 1/1 \end{bmatrix} \approx \begin{bmatrix} 1.0000 & 0.3333 & 4.0000 \\ 3.0000 & 1.0000 & 5.0000 \\ 0.2500 & 0.2000 & 1.0000 \end{bmatrix}$$

where the element m_{ij} of the matrix $\mathbf{M}$ indicates the relative importance of criteria Ci with respect to the criteria Cj. In this example, $m_{13} = 4/1$ indicates that the criteria C1 (Fuel economy) is four times as important as the criteria C3 (Style) for the DM, etc. From this pairwise matrix, Saaty demonstrated that the ranking of the priorities of the criteria can be obtained from the normalized eigenvector[3], denoted $\mathbf{w}$, associated with the principal eigenvalue of the matrix, denoted λ. In this example, one has $\lambda = 3.0857$ and $\mathbf{w} = [0.2797\ 0.6267\ 0.0936]'$ which shows that C2 criterion (reliability) is the most important criterion with the weight 0.6267, then the fuel economy criterion C1 is the second most important criterion with weight 0.2797, and finally C3 criterion (Style) is the least important criterion with weight 0.0936 for the DM. A similar ranking procedure can be used to find the relative weights of each car A, B, C or D with respect to each criterion C1, C2 and C3 based on given DM preferences, hence one will get three new normalized eigenvectors denoted $\mathbf{w}(C1)$, $\mathbf{w}(C2)$ and $\mathbf{w}(C3)$. By example, if one has the following normalized vectors

$$[\mathbf{w}(C1)\,\mathbf{w}(C2)\,\mathbf{w}(C3)] = \begin{bmatrix} 0.2500 & 0.4733 & 0.1129 \\ 0.1304 & 0.0611 & 0.4435 \\ 0.5109 & 0.1832 & 0.0565 \\ 0.1087 & 0.2824 & 0.3871 \end{bmatrix}$$

then the solution of the MCDM problem (here the selection of the "best" car according to the DM multicriteria preferences) is finally obtained by multiplying the matrix $[\mathbf{w}(C1)\ \mathbf{w}(C2)\ \mathbf{w}(C3]$ by the criteria ranking vector $\mathbf{w}$. For this example, one will get:

$$\begin{bmatrix} 0.2500 & 0.4733 & 0.1129 \\ 0.1304 & 0.0611 & 0.4435 \\ 0.5109 & 0.1832 & 0.0565 \\ 0.1087 & 0.2824 & 0.3871 \end{bmatrix} \times \begin{bmatrix} 0.2797 \\ 0.6267 \\ 0.0936 \end{bmatrix} = \begin{bmatrix} 0.3771 \\ 0.1163 \\ 0.2630 \\ 0.2436 \end{bmatrix}$$

Based on this result, the car A which has the most important weight (0.3771) will be selected by the DM. The costs could also be included in AHP by taking into account the benefit to cost ratios which will allow to chose alternative with lowest cost and highest benefit. For example, let's suppose that the cost of car A is 21000 euros, the cost of car B is 13000 euros, the cost of car C is 12000 euros and the cost of car D is 18000 euros, then the normalized cost vector is $[0.3281\ 0.2031\ 0.1875\ 0.2812]'$, so that the benefit-cost ratios are now $[0.3771/0.3281 = 1.1492\ 0.1163/0.2031 = 0.5724\ 0.2630/0.1875 = 1.4026\ 0.2436/0.2812 = 0.8663]'$. Taking into account now the cost of vehicles, now the best solution for the DM is to choose the car C since it offers the highest benefit-cost ratio.

In this paper we do not focus on the rank reversal problem of AHP as discussed in [9], [10], [13], [18], [22], but we propose an extension of AHP using aggregation method developed in DSmT framework, able to make a difference between importance of criteria, uncertainty related to the evaluations of criteria and reliability of the different sources.

[2]The relationships between preferences given by a DM may not be transitive as shown in this example, nevertheless one has to deal with these inputs even in such situations.

[3]Note that if the relationships on the criteria is transitive, then we can easily construct the normalized vector of priorities from a system of algebraic equations, without employing Saaty's matrix approach. For example if in the previous example one assumes[4] $M_{23} = 12/1$ and $M_{32} = 1/12$ instead of 5/1 and 1/5, then the normalized weighting vector will be directly obtained as $\mathbf{w} = [4/17\ 12/17\ 1/17]'$.

III. BASICS OF DSMT

Let $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ be a finite set of n elements assumed to be exhaustive. Θ corresponds to the frame of discernment of the problem under consideration. In general, we assume that elements of Θ are non exclusive in order to deal with vague/fuzzy and relative concepts [24], Vol. 2. This is the so-called free-DSm model. In DSmT, there is no need to work on a refined frame consisting in a discrete finite set of exclusive and exhaustive hypotheses[5] because DSm rules of combination work for any models of the frame. The hyper-power set D^Θ is defined as the set of all propositions built from elements of Θ with $\cup$ and $\cap$, see [24], Vol. 1 for examples. A (quantitative) basic belief assignment (bba) expressing the belief committed to the elements of D^Θ by a given source is a mapping $m(\cdot)$: $D^\Theta \rightarrow [0,1]$ such that: $m(\emptyset) = 0$ and $\sum_{A \in D^\Theta} m(A) = 1$. Elements $A \in D^\Theta$ having $m(A) > 0$ are called *focal elements* of $m(.)$. The credibility and plausibility functions are defined in almost[6] the same manner as in DST [23]. In DSmT, the Proportional Conflict Redistribution Rule no. 5 (PCR5) is used generally to combine bba's. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. For example: consider two bba's $m_1(.)$ and $m_2(.)$, $A \cap B = \emptyset$ for the model of Θ, and $m_1(A) = 0.6$ and $m_2(B) = 0.3$. With PCR5 the partial conflicting mass $m_1(A)m_2(B) = 0.6 \cdot 0.3 = 0.18$ is redistributed to A and B only with respect to the following proportions respectively: $x_A = 0.12$ and $x_B = 0.06$ because

$$\frac{x_A}{m_1(A)} = \frac{x_B}{m_2(B)} = \frac{m_1(A)m_2(B)}{m_1(A) + m_2(B)} = \frac{0.18}{0.9} = 0.2$$

In this paper, we work in the power set 2^Θ since most of readers are usually already familiar with this fusion space. Let's $m_1(.)$ and $m_2(.)$ be two independent[7] bba's, then the PCR5 rule is defined as follows (see [24], Vol. 2 for full justification and examples): $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1)m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)}\right] \quad (1)$$

where all denominators in (1) are different from zero. If a denominator is zero, that fraction is discarded. All propositions/sets are in a canonical form. A variant of (1), called PCR6, for combining $s > 2$ sources and for working in other fusion spaces (hyper-power sets or super power-sets) is presented in [24]. Additional properties of PCR5 can be found in [7]. Extension of PCR5 for combining qualitative bba's can be found in [24], Vol. 2 & 3.

[5]referred as Shafer's model in the literature.

[6]We just replace 2^Θ by D^Θ in the definitions of credibility and plausibility functions.

[7]i.e. each source provides its bba independently of the other sources.

IV. DSMT-AHP FOR SOLVING MCDM

DSmT-AHP aimed to perform a similar purpose as AHP [15], [16], SMART [28] or DS/AHP [2], [4], etc. that is to find the preferences rankings of the decision alternatives (DA), or groups of DA. DSmT-AHP approach consists in three steps:

- Step 1: We extend the construction of the matrix for taking into account the partial uncertainty (disjunctions) between possible alternatives. If no comparison is available between elements, then the corresponding elements in the matrix is zero. Each bba related to each (sub-) criterion is the normalized eigenvector associated with the largest eigenvalue of the "uncertain" knowledge matrix (as done in standard AHP approach).
- Step 2: We use the DSmT fusion rules, typically the PCR5 rule, to combine bba's drawn from step 1 to get a final MCDM priority ranking. This fusion step must take into account the different importances (if any) of criteria as it will be explained in the sequel.
- Step 3: Decision-making can be done based either on the maximum of belief, or on the maximum of the plausibility of Decision alternatives (DA), as well as on the maximum of the approximate subjective probability of DA obtained by different probabilistic transformations.

Example 2: Let's consider now a set of three cars $\Theta = \{A, B, C\}$ and the criteria C1=Fuel Economy, C2=Reliability. Let's assume that with respect to each criterion the following "uncertain" knowledge matrices are given:

$$\mathbf{M}(C1) = \begin{array}{c|ccc} & A & B \cup C & \Theta \\ \hline A & 1 & 0 & 1/3 \\ B \cup C & 0 & 1 & 2 \\ \Theta & 3 & 1/2 & 1 \end{array}$$

$$\mathbf{M}(C2) = \begin{array}{c|cccc} & A & B & A \cup C & B \cup C \\ \hline A & 1 & 2 & 4 & 3 \\ B & 1/2 & 1 & 1/2 & 1/5 \\ A \cup C & 1/4 & 2 & 1 & 0 \\ B \cup C & 1/3 & 5 & 0 & 1 \end{array}$$

Step 1: (bba's generation) Applying AHP method, one gets the following priority vectors $\mathbf{w}(C1) \approx [0.0889\ 0.5337\ 0.3774]'$ and $\mathbf{w}(C2) \approx [0.5002\ 0.1208\ 0.1222\ 0.2568]'$ which are identified with the bba's $m_{C1}(.)$ and $m_{C2}(.)$ as follows: $m_{C1}(A) = 0.0889$, $m_{C1}(B \cup C) = 0.5337$, $m_{C1}(A \cup B \cup C) = 0.3774$ and $m_{C2}(A) = 0.5002$, $m_{C2}(B) = 0.1208$, $m_{C2}(A \cup C) = 0.1222$ and $m_{C2}(B \cup C) = 0.2568$.

Step 2: (Fusion) When the two criteria have the same full importance in the hierarchy they are fused with one of the classical fusion rules. In [4] Beynon proposed to use Dempster's rule. Here we propose to use the PCR5 fusion rule since it is known to have a better ability to deal efficiently with possibly highly conflicting sources of evidences [24], Vol. 2. With PCR5, one gets:

Elem. of 2^Θ	$m_{C1}(.)$	$m_{C2}(.)$	$m_{PCR5}(.)$
$\emptyset$	0	0	0
A	0.0889	0.5002	0.3837
B	0	0	0.1162
$A \cup B$	0	0.1208	0
C	0	0	0.0652
$A \cup C$	0	0.1222	0.0461
$B \cup C$	0.5337	0.2568	0.3887
$A \cup B \cup C$	0.3774	0	0

Step 3: (Decision-making) A final decision based on $m_{PCR5}(.)$ must be taken. Usually, the decision-maker (DM) is concerned with a single choice among the elements of Θ.

Many decision-making approaches are possible depending on the risk the DM is ready to take. A pessimistic DM will choose the singleton of Θ giving the maximum of credibility whereas an optimistic DM will choose the element having the maximum of plausibility. A fair attitude consists usually in choosing the maximum of approximate subjective probability of elements of Θ. The result however is very dependent on the probabilistic transformation (Pignistic, DSmP, Sudano's, etc) [24], Vol. 2. Below are the values of the credibility, the pignistic probability and the plausibility of A, B and C:

Elem. of Θ	$Bel(.)$	$BetP(.)$	$Pl(.)$
A	0.3837	0.4068	0.4298
B	0.1162	0.3105	0.5049
C	0.0652	0.2826	0.5000

The car A will be preferred with the pessimistic or pignistic attitudes, whereas the car B will be preferred if an optimistic attitude is adopted since one has $Pl(B) > Pl(C) > Pl(A)$.

The MCDM problem deals with several criteria having different importances and the classical fusion rules cannot be applied directly as in step 2. In AHP, the fusion is done from the product of the bba's matrix with the weighting vector of criteria. Such AHP fusion is nothing but a simple componentwise weighted average of bba's and it doesn't actually process efficiently the conflicting information between the sources. It doesn't preserve the neutrality of a full ignorant source in the fusion. To palliate these problems, we propose a solution for combining sources of different importances in the framework of DSmT and DST.

Before going further, it is essential to explain the difference between the importance and the reliability of a source of evidence. The reliability is an objective property of a source, whereas the importance of a source is a subjective characteristic expressed by the fusion system designer. The reliability of a source represents its ability to provide the correct assessment/solution of the given problem. It is characterized by a discounting reliability factor, usually denoted α in $[0,1]$, which should be estimated from statistics when available, or by other techniques [11]. The reliability can be context-dependent. By convention, we usually take $\alpha = 1$ when the source is fully reliable and $\alpha = 0$ if the source is totally unreliable. The reliability of a source is usually taken into account with Shafer's discounting method [23] defined by:

$$\begin{cases} m_\alpha(X) = \alpha \cdot m(X), \quad \text{for } X \neq \Theta \\ m_\alpha(\Theta) = \alpha \cdot m(\Theta) + (1-\alpha) \end{cases} \tag{2}$$

The importance of a source is not the same as its reliability and it can be characterized by an importance factor, denoted β in $[0,1]$ which represents somehow the weight of importance granted to the source by the fusion system designer. The choice of β is usually not related with the reliability of the source and can be chosen to any value in $[0,1]$ by the designer for his/her own reason. By convention, the fusion system designer will take $\beta = 1$ when he/she wants to grant the maximal importance of the source in the fusion process, and will take $\beta = 0$ if no importance at all is granted to this source in the fusion process. The fusion designer must be able to deal with importance factors in a different way than with reliability factors since they correspond to distinct properties associated with a source of information. The importance of a source is particularly crucial in hierarchical multi-criteria decision making problems, specially in the AHP [16], [20]. That's why it is primordial to show how the importance can be efficiently managed in evidential reasoning approaches. The main question we are concerned here is how to deal with different importances of sources in the fusion process in such a way that a clear distinction is made/preserved between reliability and importance? Our preliminary investigations for the search of the solution of this problem were based on the self/auto-combination of the sources. But such approach is very disputable and cannot be used satisfactorily in practice whatever the fusion rule is adopted because it can be easily shown that the auto-conflict tends quickly to 1 after several auto-fusions [11]. Actually a better approach can be used for taking into account the importances of the sources and can be considered as the dual of Shafer's discounting approach for reliabilities of sources. The idea was originally introduced briefly by Tacnet in [24], Vol.3, Chap. 23, p. 613. It consists to define the importance discounting with respect to the empty set rather than the total ignorance Θ (as done with Shafer's discounting). Such new discounting deals easily with sources of different importances and is very simple to use. Mathematically, we define the importance discounting of a source $m(.)$ having the importance factor β in $[0,1]$ by:

$$\begin{cases} m_\beta(X) = \beta \cdot m(X), \quad \text{for } X \neq \emptyset \\ m_\beta(\emptyset) = \beta \cdot m(\emptyset) + (1-\beta) \end{cases} \tag{3}$$

Here we allow to deal with non-normal bba since $m_\beta(\emptyset) \geq 0$ as suggested by Smets in [26]. This new discounting preserves the specificity of the primary information since all focal elements are discounted with same importance factor. Here we use the positive mass of the empty set as an intermediate/preliminary step of the fusion process. Clearly when $\beta = 1$ is chosen by the fusion designer, it will mean that the source must take its full importance in the fusion process and so the original bba $m(.)$ is kept unchanged. If the fusion designer takes $\beta = 0$, one will deal with $m_\beta(\emptyset) = 1$ which is interpreted as a fully non important source. $m(\emptyset) > 0$ is not interpreted as the mass committed to some conflicting information (classical interpretation), nor as the mass committed to unknown elements when working with the open-world assumption (Smets interpretation), but only as the mass of the discounted importance of a source in this particular context. Based on this discounting, one adapts PCR5 (or PCR6) rule for $N \geq 2$ discounted bba's $m_{\beta,i}(.)$, $i = 1, 2, \ldots N$ by considering the following extension, denoted PCR5$_\emptyset$, defined by: $\forall X \in 2^\Theta$

$$m_{PCR5_\emptyset}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)} \right] \tag{4}$$

A similar extension can be done for PCR5 and PCR6 formulas for $N > 2$ sources given in [24], Vol. 2. A detailed presentation of this technique with several examples will appear in [25] and thus it is not reported here. The difference between eqs. (1) and (4) is that $m_{PCR5}(\emptyset) = 0$ whereas $m_{PCR5_\emptyset}(\emptyset) \geq 0$. Since we usually work with normal bba's for decision making support, the combined bba will be normalized. In the AHP context, the importance factors correspond to the components of the normalized eigenvector $\mathbf{w}$.

Example 3: Take back example 2 assume that C2 (the reliability) is three times more important than C1 (fuel economy) so that the knowledge matrix is given by:

$$\mathbf{M} = \begin{bmatrix} 1/1 & 1/3 \\ 3/1 & 1/1 \end{bmatrix} \approx \begin{bmatrix} 1.0000 & 0.3333 \\ 3.0000 & 1.0000 \end{bmatrix}$$

Its normalized principal eigenvector is $\mathbf{w} = [0.2500\ 0.7500]'$ and indicates that C2 is three times more important than C1 as expressed in the prior DM preferences for ranking criteria. $\mathbf{w} = [w_1\ w_2]'$ can also be obtained directly by solving the algebraic system of equations $w_2 = 3w_1$ and $w_1 + w_2 = 1$ with $w_1, w_2 \in [0, 1]$. If we apply the importance discounting with $\beta_1 = w_1 = 0.25$ and $\beta_2 = w_2 = 0.75$, one gets the following discounted bba's

Elem. of 2^Θ	$m_{\beta_1,C1}(.)$	$m_{\beta_2,C2}(.)$
$\emptyset$	0.7500	0.2500
A	0.0222	0.3751
B	0	0
$A \cup B$	0	0.0906
C	0	0
$A \cup C$	0	0.0917
$B \cup C$	0.1334	0.1926
$A \cup B \cup C$	0.0944	0

With the PCR5$_\emptyset$ fusion of the sources $m_{\beta_1,C1}(.)$ and $m_{\beta_2,C2}(.)$, one gets the results in the table. For decision-making support, one prefers to work with normal bba's. Therefore $m_{PCR5_\emptyset}(.)$ is normalized by redistributing back $m_{PCR5_\emptyset}(\emptyset)$ proportionally to the masses of other focal elements as shown in the right column of the next table.

Elem. of 2^Θ	$m_{PCR5_\emptyset}(.)$	$m^{\text{normalized}}_{PCR5_\emptyset}(.)$
$\emptyset$	0.6558	0
A	0.1794	0.5213
B	0.0121	0.0351
$A \cup B$	0.0159	0.0461
C	0.0122	0.0355
$A \cup C$	0.0161	0.0469
$B \cup C$	0.1020	0.2963
$A \cup B \cup C$	0.0065	0.0188

If all sources have the same full importances (i.e. all β_i=1), then $m_{PCR5_\emptyset}(.) = m_{PCR5}(.)$ which is normal because in such case $m_{\beta_i=1,Ci}(.) = m_{Ci}(.)$. From $m^{\text{normalized}}_{PCR5_\emptyset}(.)$ one can easily compute the credibility, pignistic probability or plausibility of each element of Θ for decision-making. In this example one gets:

Elem. of Θ	$Bel(.)$	$BetP(.)$	$Pl(.)$
A	0.5213	0.5741	0.6331
B	0.0351	0.2126	0.3963
C	0.0355	0.2134	0.3974

If the classical AHP "fusion" method (i.e. weighted arithmetic mean) is used directly with bba's $m_{C1}(.)$ and $m_{C2}(.)$, one gets:

$$m_{AHP}(.) = \begin{bmatrix} 0 & 0 \\ 0.0889 & 0.5002 \\ 0 & 0 \\ 0 & 0.1208 \\ 0 & 0 \\ 0 & 0.1222 \\ 0.5337 & 0.2568 \\ 0.3774 & 0 \end{bmatrix} \times \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.3974 \\ 0 \\ 0.0906 \\ 0 \\ 0.0917 \\ 0.3260 \\ 0.0944 \end{bmatrix}$$

which would have provided the following result for decision-making

Elem. of Θ	$Bel(.)$	$BetP(.)$	$Pl(.)$
A	0.3974	0.5200	0.6741
B	0	0.2398	0.5110
C	0	0.2403	0.5121

In this very simple example, one sees that the importance discounting technique coupled with PCR5-based fusion rule (what we call the DSmT-AHP approach) will suggest, as with classical AHP, to choose the alternative A since the car A has a bigger credibility (as well as a bigger pignistic probability and plausibility) than cars B or C. It is however worth to note that the values of $Bel(.)$, $BetP(.)$ and $Pl(.)$ obtained by both methods are slightly different. The difference in results can have a strong impact in practice in the final result for example if the costs of vehicles have also to be included in the final decision (as explained at the end of the example 1). Note also that the uncertainties $U(X) = Pl(X) - Bel(X)$ of alternatives $X = A, B, C$ have been seriously diminished when using DSmT-AHP with respect to what we obtain with classical AHP as seen in the following table. The uncertainty reduction is a nice expected property specially important for decision-making support.

Elem. of Θ	$U(.)$ with AHP	$U(.)$ with DSmT-AHP
A	0.2767	0.1118
B	0.5110	0.3612
C	0.5121	0.3619

Important remark: If Dempster's rule is used instead of PCR5$_\emptyset$ rule, one gets the following results when comparing the fusion of $m_{C1}(.)$ with $m_{C2}(.)$ (i.e. without importance discounting) with the fusion of $m_{\beta_1=w_1=0.25,C1}(.)$ with $m_{\beta_2=w_2=0.75,C2}(.)$ (i.e. with importance discounting of criteria C1 and C2):

Elem. of 2^Θ	$m_{DS}(.)$	$m_{DS,\mathbf{w}}(.)$
$\emptyset$	0	0
A	0.3588	0.3588
B	0.0908	0.0908
$A \cup B$	0.0642	0.0642
C	0.0918	0.0918
$A \cup C$	0.0649	0.0650
$B \cup C$	0.3294	0.3294
$A \cup B \cup C$	0	0

Clearly, Dempster's rule cannot deal properly with importance discounted bba's as we have proposed in this work just because the importance discounting technique preserves the specificity of the primary information and thus Dempster's rule does not make a difference in results when combining either $m_{C1}(.)$ with $m_{C2}(.)$ or when combining $m_{\beta_1\neq1,C1}(.)$ with $m_{\beta_2\neq1,C2}(.)$ due to the way of processing of the total conflicting mass of belief. PRC5 deals more efficiently with importance discounted bba's as we have shown in this example. So it is not surprising that such discounting technique has never been proposed and used in DST framework and this explains why only the classical Shafer's discounting technique (the reliability discounting) is generally adopted. By using Dempster's rule, the fusion designer has no other choice but to consider importance and reliability as same notions ! The DSmT framework with PCR5 (or PCR6) rule and the importance discounting technique proposed here provides an interesting and simple solution for the fusion of sources with different importances which makes a clear distinction between importances and reliabilities of sources.

V. DSmT-AHP for solving MCGDM

Previously, a new approach mixing AHP with DSmT solving MCDM problem has been presented. In many practical situations however, the decision must be taken by a group of $n > 1$ Decision Makers (GDM), denoted $GDM = \{DM_i, i = 1, 2, \ldots, n\}$, rather than a single DM, and from the Multi-Criteria preference rankings of the DM_i's. The importance (influence) of each member of the GDM is usually non-equivalent [1] and the importance of each DM of the GDM must be efficiently taken into account in the final decision-making process. Let's denote by $m_{DM_i}(.)$ the result of DSmT-AHP approach (see section IV) related with $DM_i \in GDM$. The MCGDM problem consists in combining all opinions/preferences rankings $m_{DM_i}(.)$, $i = 1, \ldots, n$ with their own (possibly different) importances. When all DM_i's have equal importance, the classical fusion rules[8] $\oplus$ for combining $m_{DM_i}(.)$ can be directly used to get the final result $m_{MCGDM}(.) = [m_{DM_1} \oplus m_{DM_2} \oplus \ldots \oplus m_{DM_n}](.)$; If the DM_i's have different importance weights w_i, the DSmT-AHP approach can also be used at the GDM fusion level using the importance discounting approach presented here. The result for group decision-making is given by the PCR5$_\emptyset$ fusion of $m_{\beta_i, DM_i}(.)$, with $\beta_i = w_i$ and then the result must be normalized for decision making support. In [6], Beynon used the classical discounting technique [23] to readjust $m_{DM_i}(.)$ with w_i's and he identified the importance factors with the reliability factors. In our opinions, this is disputable since importance of a DM_i is not necessarily related with its reliability but rather with the importance in the problem of the choice of his/her Multi-Criteria to establish his/her ranking, or it can come from other (political, hierarchical, etc.) reasons. In our new approach, we make a clear distinction between notions of importance and reliability and both notions can be easily taken into account [25] with DSmT-AHP for solving MCGDM problems, i.e. we can use the classical discounting technique for taking into the reliabilities of the sources, and use the importance discounting proposed here for dealing with the importances of sources.

VI. Conclusions and perspectives

In this paper, we have presented a new method for Multi-Criteria Decision-Making (MCDM) and Multi-Criteria Group Decision-Making (MCGDM) based on the combination of AHP method developed by Saaty and DSmT. The AHP method allows to build bba's from DM preferences of solutions which are established with respect to several criteria. The DSmT allows to aggregate efficiently the (possibly highly conflicting) bba's based on each criterion. This DSmT-AHP method allows to take into account also the different importances of the criteria and/or of the different members of the decision-makers group. The application of this DSmT-AHP approach for the prevention of natural hazards in mountains is currently under progress, see [24], Vol.3, Chap. 23, and [27].

[8] typically the PCR5 or PCR6 rules, or eventually Dempster's rule if the conflict between DM_i's is low.

References

[1] J. Barzilai, F.A. Lootsma, *Power relations and group aggregation in the multiplicative AHP and SMART*, J. of MCDA, Vol. 6, pp. 155–165, 1997.

[2] M. Beynon, B. Curry, P.H. Morgan, *The Dempster-Shafer theory of evidence: An alternative approach to multicriteria decision modelling*, Omega, Vol. 28, No. 1, pp. 37–50, 2000.

[3] M. Beynon, D. Cosker, D. Marshall, *An expert system for multi-criteria decision making using Dempster-Shafer theory*, Expert Syst. with Appli., Vol. 20, No. 4, pp. 357–367, 2001.

[4] M. Beynon, *DS/AHP method: A mathematical analysis, including an understanding of uncertainty*, Eur. J. of Oper. Res., Vol. 140, pp. 148–164, 2002.

[5] M. Beynon, *Understanding local ignorance and non-specificity within the DS/AHP method of multi-criteria decision making*, Eur. J. of Oper. Res., Vol. 163, pp. 403–417, 2005.

[6] M. Beynon, *A method of aggregation in DS/AHP for group decision-making with non-equivalent importance of individuals in the group*, Comp. & Oper. Res., No. 32, pp. 1881-1896, 2005.

[7] J. Dezert, F. Smarandache, *Non Bayesian conditioning and deconditioning*, International Workshop on Belief Functions, Brest, France, April 2010.

[8] E.H. Forman, S.I. Gass, *The analytical hierarchy process: an exposition*, Oper. Res., Vol. 49, No. 4 pp. 46–487, 2001.

[9] R.D. Holder, *Some Comment on the Analytic Hierarchy Process*, J. of the Oper. Res. Soc., Vol. 41, No. 11, pp. 1073–1076, 1990.

[10] F.A. Lootsma, *Scale sensitivity in the multiplicative AHP and SMART*, J. of MCDA, Vol. 2, pp. 87–110, 1993.

[11] A. Martin, A.-L. Jousselme, C. Osswald, *Conflict measure for the discounting operation on belief functions*, Proc. of Fusion 2008, Cologne, Germany, July 2008.

[12] C. K. Murphy, *Combining belief functions when evidence conflicts*, Dec. Sup. Syst., vol. 29, pp. 1-9, 2000.

[13] J. Perez, *Some comments on Saaty's AHP*, Manag. Sci., Vol. 41, No. 6, pp. 1091–1095, 1995.

[14] J. Perez, J.L. Jimeno, E. Mokotoff, *Another Potential Shortcoming of AHP*, TOP: An Official J. of the Spanish Soc. of Stats and Oper. Res., Vol. 14, No. 1 , June, 2006.

[15] T.L. Saaty, *A scaling method for priorities in hierarchical structures*, J. of Math. Psych., Vol. 15, PP. 59–62, 1977.

[16] T.L. Saaty, *The Analytical Hierarchy Process*, McGraw Hill, New York, 1980.

[17] T.L. Saaty, *An exposition of the AHP in reply to the paper "Remarks on the Analytic Hierarchy Process"*, Manag. Sci., Vol. 36, No. 3, pp. 259–268, 1990.

[18] T.L. Saaty, *Response to Holder's Comments on the Analytic Hierarchy Process*, J. of the Oper. Res. Soc., Vol. 42, No. 10, pp. 909–914, October 1991.

[19] T.L. Saaty, *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World*, RWS Publ., Pittsburgh, PA, USA, 1999.

[20] T.L. Saaty, *Fundamentals of decision making and priority theory with the analytic hierarchy process*, Vol. VI of the AHP series, RWL Publ., Pittsburgh, PA, USA.

[21] T.L. Saaty, *Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors - The Analytic Hierarchy/Network Process*, RACSAM, Vol. 102, No. 2, pp. 251–318, 2008.

[22] S. Schenkerman, *Inducement of nonexistent order by the analytic hierarchy process*, Dec. Sci., Spring 1997.

[23] G. Shafer, *A mathematical theory of evidence*, Princeton Univ. Press, 1976.

[24] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vol. 1-3, ARP, 2004–2009. http://www.gallup.unm.edu/~smarandache/DSmT.htm

[25] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, submitted March 2010.

[26] Ph. Smets Ph., *The Combination of Evidence in the Transferable Belief Model*, IEEE Trans. PAMI 12, pp. 447–458, 1990.

[27] J.-M. Tacnet, M. Batton-Hubert, J. Dezert, *A two-step fusion process for multi-criteria decision applied to natural hazards in mountains*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[28] D. Von Winterfeldt, W. Edwards *Decision analysis and behavioral research*, Cambridge Univ. Press, USA, 1986.

Sequential Adaptive Combination of Unreliable Sources of Evidence

Zhun-ga Liu
Quan Pan
Yong-mei Cheng
Jean Dezert

Originally published as Liu Z., Pan Q., Cheng Y., Dezert J., *Sequential adaptive combination of unreliable sources of evidence*, In proc. of International Workshop on Belief Functions, Brest, France, April 2-4, 2010, and reprinted with permission.

***Abstract*-In theories of evidence, several methods have been proposed to combine a group of basic belief assignments altogether at a given time. However, in some applications in defense or in robotics the evidences from different sources are acquired only sequentially and must be processed in real-time and the combination result needs to be updated the most recent information. An approach for combining sequentially unreliable sources of evidence is presented in this paper. The sources of evidence are not considered as equi-reliable in the combination process, and no prior knowledge on their reliability is required. The reliability of each source is evaluated on the fly by a distance measure, which characterizes the variation between one source of evidence with respect to the others. If the source is considered as unreliable, then its evidence is discounted before entering in the fusion process. Dempster's rule of combination and its main alternatives including Yager's rule, Dubois and Prade rule, and PCR5 are adapted to work under different conditions. In this paper, we propose to select the most adapted combination rule according to the value of conflicting belief before combining the evidence. The last part of this paper is devoted to a numerical example to illustrate the interest of this approach.**

Keywords: evidence theory, combination rule, evidence distance, conflicting belief.

I. Introduction

Evidence theories[1] are widely applied in the field of information fusion. A particular attention has been focused on how to efficiently combine sources of evidence altogether at the same time (static approach), and many rules aside Dempster's rule have been proposed [1], [2], [6], [9]. In many applications however, the evidences from different sources are acquired sequentially by different sensors or human experts and the belief updating and decision-making need to be taken in real-time which requires a sequential/dynamic approach rather than a static approach of the fusion problem.

Usually the evidences arising from different independent sources are often considered equally reliable in the combination process, when the prior knowledge about the reliability of each source is unknown. However, all the sources of evidence to be combined can have different reliabilities in real applications. If the sources of evidence are considered as equi-reliable, the unreliable ones may bring a very bad influence on combination result, and even leads to inconsistent results and wrong decisions. Thus, the reliability of each source must be taken account in the fusion process as best as possible to provide a useful and unbiased result. In this work, we propose to evaluate on the fly the relability of the sources to combine based on an evidential distance/reliability measure. From this reliability measure, one can discount accordingly the unreliable sources before applying a rule of combination of basic belief assignments (bba's).

Many rules, like Dempster's rule [7] and its alternatives can be used to combine sources of evidences expressed by bba's and they all have their drawbacks and advantages (see [8], Vol. 1, for a detailed presentation). Dempster's rule, is usually considered well adapted for combining the evidences in low conflict situations and it requires acceptable complexity when the dimension of the frame of discernment is not too large. Dempster's rule however provides counter-intuitive behaviors when the sources evidences become highly conflicting. To palliate this drawback, several interesting alternatives have been proposed when Dempster's rule doesn't work well, mainly: Yager's rule [9], Dubois and Prade rule (DP rule) [2], and PCR5 (proportional conflict redistribution rule no 5) [8] developed in DSmT framework. The difference among Dempster's rule and its main alternatives mainly lies in the distribution of the conflicting belief $m_{\oplus}(\emptyset)$ which is generally used to characterize the total amount of conflict [4] between sources. In this paper, we propose to select the proper rule of combination based on the value of the total degree of conflict $m_{\oplus}(\emptyset)$. The last part of this paper presents a numerical example to show how the approach of sequential adaptive combination of unreliable sources of evidence works.

II. Preliminaries

A. Basics of Dempster-Shafer theory (DST)

DST [7] is developed in Shafer's model. In this model, a fixed set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ is called the frame of discernment of fusion problem. All the elements in Θ are mutually exclusive and exhaustive. The set of all subsets of

[1]DST (Dempster-Shafer Theory) [7] or DSmT (Dezert-Smarandache Theory) [8].

Θ is called the power set of Θ, and it is denoted 2^Θ. For instance, if $\Theta = \{\theta_1, \theta_2, \theta_3\}$, then $2^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_3, \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cup \theta_2 \cup \theta_3\}$. A basic belief assignment (bba), also called mass of belief, is a mapping $m : 2^\Theta \to [0, 1]$ associated to a given body of evidence $\mathcal{B}$ such that $m(\emptyset) = 0$ and $\sum_{A \in 2^\Theta} m(A) = 1$. The credibility (also called belief) of $A \subseteq \Theta$ is defined by $Bel(A) = \sum_{\substack{B \in 2^\Theta \\ B \subseteq A}} m(B)$. The commonality function $q(.)$ and the plausibility function $Pl(.)$ are also defined by Shafer in [7]. The functions $m(.)$, $Bel(.)$, $q(.)$ and $Pl(.)$ are in one-to-one correspondence.

Let $m_1(.)$ and $m_2(.)$ be two bba's provided by two independent bodies of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$ over the frame of discernment Θ. The fusion/combination of $m_1(.)$ with $m_2(.)$, denoted $m(.) = [m_1 \oplus m_2](.)$ is obtained in DST with Dempster's rule of combination as follows:

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \frac{\sum_{X_1 \cap X_2 = A} m_1(X_1) m_2(X_2)}{\sum_{X_1 \cap X_2 \neq \emptyset} m_1(X_1) m_2(X_2)} \quad \forall A \neq \emptyset, A \in 2^\Theta \end{cases} \tag{1}$$

The degree of conflict between the bodies of evidence $\mathcal{B}_1$ and $\mathcal{B}_2$ is defined by

$$m_\oplus(\emptyset) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2) \tag{2}$$

Dempster's rule can be directly extended to the combination of S independent and equally reliable sources. It is a commutative and associative rule of combination and it preserves the neutral impact of the vacuous belief assignment defined by $m_{vba}(\Theta) = 1$.

B. Main alternatives to Dempster's rule

Dempster's rule yields counterintuitive results when the evidences highly conflict because of its way of assigning the mass of conflicting belief $m_\oplus(\emptyset)$. Thus, a lot of alternatives to Dempster's rule have been proposed for overcoming limitations of Dempster's rule. The main alternative rules including Yager's rule [9], DP rule [2] and PCR5 [8] are briefly recalled.

- **Yager's rule**: Yager admits the conflicting belief is not reliable. So $m_\oplus(\emptyset)$ is transferred to the total ignorance in Yager's rule. It is given by $m(\emptyset) = 0$ and for $A \neq \emptyset$, $A \in 2^\Theta$ by

$$\begin{cases} m(A) = \sum_{\substack{X, Y \in 2^\Theta \\ X \cap Y = A}} m_1(X) m_2(Y), \text{ for } A \neq \Theta \\ m(\Theta) = m_1(\Theta) m_2(\Theta) + \sum_{\substack{X, Y \in 2^\Theta \\ X \cap Y = \emptyset}} m_1(X) m_2(Y) \end{cases} \tag{3}$$

- **Dubois & Prade rule**: This rule assumes that if two sources of evidence are in conflict, one of them is right but we don't know which one. Thus, if $X \cap Y = \emptyset$, then the mass committed to the set $X \cap Y$ by the conjunctive operator should be transferred to $X \cup Y$. According to this principle, DP rule is defined by $m(\emptyset) = 0$ and for $A \neq \emptyset$ and $A \in 2^\Theta$ by

$$m(A) = \sum_{\substack{X, Y \in 2^\Theta \\ X \cap Y = A}} m_1(X) m_2(Y) + \sum_{\substack{X, Y \in 2^\Theta \\ X \cap Y = \emptyset \\ X \cup Y = A}} m_1(X) m_2(Y) \tag{4}$$

- **PCR5 rule**: PCR5 transfers the partial conflicting mass to the elements involved in the conflict, and it is considered as the most mathematically exact redistribution of conflicting mass to nonempty sets following the logic of the conjunctive rule. PCR5 is defined by $m(\emptyset) = 0$ and $\forall A \neq \emptyset$, $A \in 2^\Theta$ by

$$m(A) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = A}} m_1(X_1) m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap A = \emptyset}} \left[\frac{m_1(A)^2 m_2(X_2)}{m_1(A) + m_2(X_2)} + \frac{m_2(A)^2 m_1(X_2)}{m_2(A) + m_1(X_2)} \right] \tag{5}$$

The details, examples and the extension of PCR5 formula (5) for $S > 2$ sources are given in [8].

C. Discounting source of evidence

When the sources of evidences are not considered equally reliable, it is reasonable to discount each unreliable source s_i, $i = 1, 2, \ldots, S$ by a reliability factor $\alpha_i \in [0, 1]$. Following the classical discounting method [7], a new discounted bba $m'(.)$ is obtained from the initial bba $m(.)$ provided by the unreliable source s_i as follows

$$\begin{cases} m'(A) = \alpha_i \cdot m(A), \quad A \neq \Theta \\ m'(\Theta) = 1 - \sum_{\substack{A \in 2^\Theta \\ A \neq \Theta}} m'(A) \end{cases} \tag{6}$$

$\alpha_i = 1$ means that the total confidence in the source s_i, and the original bba doesn't need to be discounted. $\alpha_i = 0$ means that the source is s_i is totally unreliable and its bba is revised as a vacouous bba $m'(\Theta) = 1$, which will have a neutral impact in the fusion process. In practice, the discounting method can be used efficiently if one has a good estimation of the reliability factor of each source. We show in the next section how one can evaluate the relability of a source.

III. Evaluating the reliability of each source

Without prior knowledge on the reliability of the sources of evidence, we propose to evaluate the reliability factors of each source based on the distance between the bba from a given source s_i with respect to the others. If the bba of the given source, say s_i varies too much with respect to the others, this source of evidence is considered not reliable and it will be discounted before to be combined. We will show further how the discounting/reliability factor can be estimated. We implicitly assume here that the following principle "Truth is reflected by the majority of opinions" holds.

In [3], Jousselme et al. have proposed the following distance measure $d_J(\mathbf{m}_1, \mathbf{m}_2)$ between two bba's[2] $\mathbf{m}_1 \triangleq m_1(.)$ and $\mathbf{m}_2 \triangleq m_2(.)$ defined on the same power set 2^Θ:

$$d_J(\mathbf{m}_1, \mathbf{m}_2) = \sqrt{\frac{1}{2}(\mathbf{m}_1 - \mathbf{m}_2)^T \mathbf{D}(\mathbf{m}_1 - \mathbf{m}_2)} \quad (7)$$

where $\mathbf{D}$ is a $2^{|\Theta|} \times 2^{|\Theta|}$ positive matrix whose elements are defined as $D_{ij} \triangleq \frac{|A_i \cap B_j|}{|A_i \cup B_j|}$ where A_i and B_j are elements of the power set 2^Θ. $d_J(\mathbf{m}_1, \mathbf{m}_2) \in [0, 1]$ is a distance which measures the similarity between $\mathbf{m}_1$ and $\mathbf{m}_2$ considering both the values and the relative specificity of focal elements of each bba.

The total degree of conflict $m_\oplus(\emptyset)$ obtained from all focal elements which are incompatible doesn't actually capture the similarity between bba's as shown by Martin et al. in [5].

If N pieces of evidence $\mathbf{m}_1$, $\mathbf{m}_2$, ..., $\mathbf{m}_N$ are combined sequentially, two approaches similarly with [5] could be used to measure the variation between $\mathbf{m}_j$ and the others. One considers the average value d_J between $\mathbf{m}_j$ and the others which is given by

$$d_1^{j-1}(\mathbf{m}_j) = \frac{1}{j-1} \sum_{i=1}^{j-1} d_J(\mathbf{m}_j, \mathbf{m}_{j-i}) \quad (8)$$

The other one is simply defined as

$$d_2^{j-1}(\mathbf{m}_j) = d_J(\mathbf{m}_j, \mathbf{m}_1^{j-1}) \quad (9)$$

where $\mathbf{m}_1^{j-1} \triangleq m_1^{j-1}(.)$ is obtained by the sequential combinination of the bba's $m_1(.)$, $m_2(.)$,..., $m_{j-1}(.)$, i.e. $m_1^{j-1}(.) = (((m_1 \oplus m_2) \oplus m_3) \cdots \oplus m_{j-1})(.)$ with a fusion rule such as Dempster's rule, Yager's rule, DP rule, PCR5, etc. The second measure, $d_2^{j-1}(\mathbf{m}_j)$, reflects only the difference between $\mathbf{m}_j$ and the combined bba $\mathbf{m}_1^{j-1}$ and thus cannot precisely measure the similarity between m_j and the other individual evidences $m_1(.)$, $m_2(.)$, ..., $m_{j-1}(.)$ because some information on specificities of these individual bba's has been lost forever through the fusion process. The following examples will show the distinction between these two methods.

Example 1: Let's consider the frame of discernment $\Theta = \{A, B, C\}$, Shafer's model and the same following bba's

$$\begin{aligned}
m_1(.): &\quad m_1(A) = 0.5, m_1(B) = 0.2 \\
&\quad m_1(A \cup B) = m_1(C) = m_1(\Theta) = 0.1 \\
m_2(.): &\quad m_2(A) = 0.5, m_2(B) = 0.2 \\
&\quad m_2(A \cup B) = m_2(C) = m_2(\Theta) = 0.1 \\
&\vdots \\
m_j(.): &\quad m_j(A) = 0.5, m_j(B) = 0.2 \\
&\quad m_j(A \cup B) = m_j(C) = m_j(\Theta) = 0.1
\end{aligned}$$

The difference between $m_j(.)$, for $j \geq 2$, and all the bba's $m_i(.)$, for $i < j$ according to formula (8) gives $d_1^{j-1}(\mathbf{m}_j) = 0$, which shows correctly that $m_j(.)$ is identical to the other bba's $m_i(.)$, for $i < j$. If one uses the measure $d_2^{j-1}(\mathbf{m}_j)$ defined in (9), one gets the results plotted in Fig. 1.

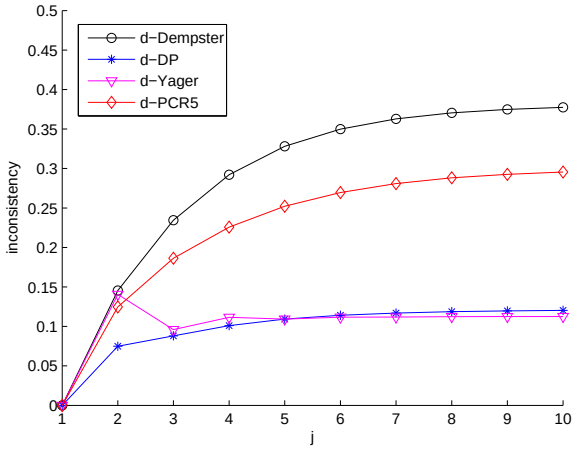


Fig. 1: Variation of the similarity measure $d_2^{j-1}(\mathbf{m}_j)$ based on different fusion rules.

One sees that there exists a variation of the similarity measure using all different fusion rules with a trend to certain values when j increases. This is a bad behavior since we know that $\mathbf{m}_j$ equals to the others bba's and we would expect to get $d_2^{j-1}(\mathbf{m}_j) = 0$ which unfortunately is not the case. That is the main reason why we abondon the use of $d_2^{j-1}(\mathbf{m}_j)$ measure in the sequel of this work.

Example 2: Let's consider the frame of discernment $\Theta = \{A, B, C, D, E\}$, Shafer's model, and the following bba's

$$\begin{aligned}
m_1(.): &\quad m_1(A) = 0.6, m_1(B) = m_1(C) = 0.1 \\
&\quad m_1(D) = m_1(E) = 0.1 \\
m_2(.): &\quad m_2(\Theta) = 1 \\
m_3(.): &\quad m_3(\Theta) = 1 \\
&\vdots \\
m_{j-1}(.): &\quad m_{j-1}(\Theta) = 1 \\
m_j(.): &\quad m_j(A) = 0.6, m_j(B) = m_j(C) = 1, \\
&\quad m_j(D) = m_j(E) = 0.1
\end{aligned}$$

In this example, $m_j(.) = m_1(.)$, but $m_j(.)$ is quite different from the others bba's $m_i(.)$, $i \neq 1$. The similarity measure $d_1^{j-1}(\mathbf{m}_j)$ between $m_j(.)$, for $j \geq 3$ and the bba's $m_i(.)$, $i < j$ is $d_1^{j-1}(\mathbf{m}_j) = \frac{0.2(j-2)}{j-1}$ which shows a trend to 0.2 when j increases. However in such case, one always gets using Dempster's rule, Yager's rule, DP rule or PCR5 rule $d_2^{j-1}(\mathbf{m}_j) = 0$. From such very simple example, one sees that one cannot detect the dissimilarity of $\mathbf{m}_j(.)$ with a majority of quite distinct bba's when $d_2^{j-1}(\mathbf{m}_j)$ measure is used. This shows again that $d_2^{j-1}(\mathbf{m}_j)$ is actually not very appropriate for measuring the similarity between a given bba $\mathbf{m}_j(.)$ and a set of bba's. Therefore we will only consider the measure of similarity $d_1^{j-1}(\mathbf{m}_j)$ in the sequel.

[2]Here for notation convenience, we use the usual vectorial notation $\mathbf{m}_1$ and $\mathbf{m}_2$ (with boldfaced letter) for representing the entire bba's usually denoted $m_1(.)$ and $m_2(.)$. $\mathbf{m}_1$ and $\mathbf{m}_2$ are vectors of dimension $2^{|\Theta|} \times 1$. We assume that the bba's vectors are both ordered using the same order for their components.

For managing the computational burden in applications, a parameter $n \leq j-1$ is introduced in the measure $d_1^{j-1}(\mathbf{m}_j)$ and we define the new measure:

$$d_n^{j-1}(\mathbf{m}_j) = \frac{1}{\min(n, j-1)} \sum_{i=1}^{\min(n,j-1)} d_J(\mathbf{m}_j, \mathbf{m}_{j-i}) \quad (10)$$

The accuracy and the computational complexity of this similarity measure increases when n tends to $j-1$.

Let 's consider a given discounting tolerance threshold ω_d in $[0,1]$. If $d_n^{j-1}(\mathbf{m}_j) \geq \omega_d$, it indicates that the bba $m_j(.)$ will be considered as not similar enough with respect to other bba's and therefore the source s_j is considered as unreliable and must be discounted before entering in the fusion process. The unreliability of the source s_j may be caused by a fault of the sensor or unexpected noises, condition changes, etc. In such case, the bba $m_j(.)$ needs to be discounted by formula (6). As proposed by Martin et al. in [5], the reliability factor of the source s_j is chosen as $\alpha_j = (1 - d_n^{j-1}(\mathbf{m}_j)^\lambda)^{1/\lambda}$ where the parameter λ is defined in the easiest way with $\lambda = 1$. The larger dissimilarity leads to the less reliability factor. If $d_n^{j-1}(\mathbf{m}_j) < w_d$, it means that the dissimilarity between $m_j(.)$ with other bba's is acceptable, and there is no need to revise/discount $m_j(.)$ in such case.

IV. Selection of combination rules

After evaluating the reliability of the sources, we have to select a suitable combination rule. Dempster's rule is known to offer pretty good performances when the combined bba's are not in too high conflict, otherwise when the conflict becomes too large it is generally considered safer to use alternative rules like Yager's rule, DP rule, and PCR5 rule. The following examples show the difference between the different approaches for the fusion of sources of evidences.

Example 3: This is Zadeh's example [10]. Let's consider $\Theta = \{A, B, C\}$ with Shafer's model and the following bba's

$$\begin{aligned} m_1(.): &\quad m_1(A) = 0.9, m_1(B) = 0.1 \\ m_2(.): &\quad m_2(B) = 0.1, m_2(C) = 0.9 \end{aligned}$$

One sees that the two sources are in very high conflict because the total conflict is $m_\oplus^{1,2}(\emptyset) = 0.99$. Using Dempster's rule, one gets surprisingly $m(B) = 1$ which is somehow conterintuitive since m_1 and m_2 both believe in B with a little chance, but the fusion result states that B is the only possible solution with certainty, which seems unreasonable[3]. If we use Yager's rule, DP rule, and PCR5, one gets:

- Yager's rule: $m(B) = 0.01$, $m(\Theta) = 0.99$
- DP rule: $m(B) = 0.01$, $m(A \cup B) = 0.09$, $m(A \cup C) = 0.81$, $m(B \cup C) = 0.09$
- PCR5: $m(A) = 0.486$, $m(B) = 0.028$, $m(C) = 0.4860$

These results are more reasonable in some sense, but they are not the same. Yager's rule transfers all the conflicting mass to total ignorance and produces the least specific result in the three rules. DP rule distributes the conflicting mass to the union of the involved sets, which makes the uncertainty of the result still very large. DP rule produces a less specific result than PCR5 but DP is a bit more specific than Yager's rule. PCR5 provides the most specific result since A and C share the same bba whereas B keeps a very low belief assignment.

Therefore, in order to avoid to get counterintuitive results, it is reasonable to use Yager's rule, DP rule, or PCR5 than Dempster's rule as soon as the level of conflict becomes large. The choice among Yager's rule, DP rule, and PCR5 depends on the application and the computational resource one has. PCR5 is very appropriate if a decision has to be made because it provides the most specific solution, but PCR5 requires the most computational burden. Sometimes it better to get less specific result if we don't need to take a clear/precise decision in case of high conflict between sources. In such case, Yager's rule and/or DP rule can be used instead. When the level of conflict between two bba's is low Dempster's rule can be used since it offers a good compromise between computational complexity and the specificity of the result.

Example 4: Let's consider $\Theta = \{A, B, C\}$ and

$$\begin{aligned} m_1(.): &\quad m_1(A) = 0.35, m_1(B) = 0.3, m_1(A \cup B) = 0.15, \\ &\quad m_1(C) = 0.2 \\ m_2(.): &\quad m_2(A) = 0.35, m_2(B) = 0.3, m_2(A \cup C) = 0.05, \\ &\quad m_2(A \cup B) = m_2(C) = m_2(\Theta) = 0.1 \\ m_3(.): &\quad m_3(A) = 0.3, m_3(B) = 0.3, m_3(A \cup B) = 0.2, \\ &\quad m_3(C) = m_3(A \cup C) = 0.1 \end{aligned}$$

The conflicts between each pair of bba's are given by $m_\oplus^{1,2}(\emptyset) = 0.455, m_\oplus^{1,3}(\emptyset) = 0.395, m_\oplus^{2,3}(\emptyset) = 0.395$. The levels of these conflicts are not too large according and the sequential combination $m(.) = [[m_1 \oplus m_2] \oplus m_3](.)$ using the different rules yields

- Dempster's rule: $m(A) = 0.5868$, $m(B) = 0.3592$, $m(A \cup B) = 0.0202$, $m(C) = 0.0338$
- Yager's rule: $m(A) = 0.3105$, $m(B) = 0.243$, $m(C) = 0.0555$, $m(A \cup B) = 0.097$, $m(A \cup C) = 0.0455$, $m(\Theta) = 0.2485$
- DP rule: $m(A) = 0.3255$, $m(B) = 0.2295$, $m(C) = 0.0435$, $m(A \cup B) = 0.1975$, $m(A \cup C) = 0.0575$, $m(B \cup C) = 0.0345$, $m(\Theta) = 0.112$
- PCR5: $m(A) = 0.4889$, $m(B) = 0.3941$, $m(C) = 0.0819$, $m(A \cup B) = 0.0268$, $m(A \cup C) = 0.0083$

All the rules provide reasonable results with assigning the largest belief to A, but Dempster's rule produces the most specific result with a less computational effort. Dempster's rule is thus well appropriate when $m_\oplus(\emptyset)$ is not too large.

V. Adaptive combination of sequential evidence

Here we are concerned with the real-time decision-making problem from the sequential acquisition of bba's $m_1(.)$, $m_2(.)$,..., $m_N(.)$ defined on a same frame Θ without any

[3]More generally, one can show that Dempster's rule can become insensitive to the variation of input bba's to combine - see [8], Vol. 1, Chap. 5, p. 114 for example.

prior knowledge about reliability of each source. We start with $m_1(.)$. When $m_2(.)$ is available, one combines it with $m_1(.)$ by a suitable rule according to the value of $m_{\oplus}^{1,2}(\emptyset)$ without evaluating the reliability of the two sources. When $m_j(.)$, for $j \geq 3$ becomes available at the time j, the reliability of the source s_j is evaluated and $m_j(.)$ is discounted (if necessary) by the approach presented in section III. Before combining the discounted bba $m'_j(.)$ (or $m_j(.)$ when no discounting occurs) with the last updated bba $m_1^{j-1}(.)$, the combination rule is selected according to the value of the conflict between $m_j(.)$ and $m_1^{j-1}(.)$. We use a threshold $\omega_\emptyset$. If $m_\oplus(\emptyset) < \omega_\emptyset$, Dempster's rule is selected because it offers a good compromise between complexity and specificity. Otherwise, Yager's rule, DP rule, or PCR5, are selected upon the actual application to avoid to get counterintuitive results.

The tuning of thresholds ω_d and $\omega_\emptyset$ is not easy in general. If the thresholds are too large, one takes the risk to get counterintuitive results, whereas if they are set to too low values the non specificity of the result will become large and even will lead to decision-making under big uncertainty. Therefore, both thresholds ω_d and $\omega_\emptyset$ need to be determined by accumulated experience depending on the actual application .

VI. Numerical example

Let us suppose a multisensor-based target identification system. From five independent sensors, the system collects five pieces of evidence sequentially (actually we consider here 2 possible bba's $m_{5A}(.)$ and $m_{5B}(.)$ for the fifth source). For decision-making in real-time, the combination result needs to be updated right after the new evidence arrives. The bba's defined on the power set of $\Theta = \{A, B, C\}$ are as follows

$$\begin{aligned}
m_1(.): &\quad m_1(A) = 0.8, m_1(B) = 0.1, m_1(\Theta) = 0.1\\
m_2(.): &\quad m_2(A) = 0.4, m_2(B) = 0.25,\\
&\quad m_2(C) = 0.2, m_2(B \cup C) = 0.15,\\
m_3(.): &\quad m_3(B) = 0.9, m_3(C) = 0.1,\\
m_4(.): &\quad m_4(B) = 0.45, m_4(C) = 0.45, m_4(B \cup C) = 0.1,\\
m_{5A}(.): &\quad m_{5A}(A) = 0.5, m_{5A}(A \cup B) = 0.25,\\
&\quad m_{5A}(C) = 0.1, m_{5A}(A \cup C) = 0.15,\\
m_{5B}(.): &\quad m_{5B}(B) = 0.5, m_{5B}(A \cup B) = 0.25,\\
&\quad m_{5B}(C) = 0.1, m_{5B}(B \cup C) = 0.15.
\end{aligned}$$

The five pieces of evidence are combined sequentially, and the results are presented in Table 1. The chosen thresholds are $\omega_d = 0.6, \omega_\emptyset = 0.6$ and $n = 5$.

All the rules provide reasonable results when combining consistent bba's $m_1(.)$ and $m_2(.)$. The bba $m_3(.)$ is highly conflicting with $m_1(.)$ and $m_2(.)$. If there is no prior information about the reliability of the sources, we evaluate the reliability of each source according to its variation with respect to the others. The average similarity distance between $\mathbf{m}_3$ and $\mathbf{m}_1$, $\mathbf{m}_2$ is so large that $d_n^{j-1}(\mathbf{m}_3) > \omega_d$. Thus, $m_3(.)$ is considered unreliable. If we combine directly (without discounting) $m_3(.)$ with $m_1^2(.)$ using Dempster, Yager, DP or PCR5, one gets a high belief in B with all the rules. With the adaptive rule, the bba of $m_3(.)$ is discounted with the reliability factor $\alpha = 1 - d_n^{j-1}(\mathbf{m}_3)$ to get $m'_3(.)$. The combination of $m'_3(.)$ with $m_1^2(.)$ assigns now the highest belief in A. This adaptive method is helpful to deal with the high conflicts caused by the unreliability of the sources. The difference between $m_4(.)$ and $m_1(.)$, $m_2(.)$, $m_3(.)$ is below the tolerance threshold, but the value of $m_\oplus(\emptyset)$ between $m_4(.)$ and $m_1^3(.)$ is very large, and $m_\oplus(\emptyset) = 0.8334 > \omega_\emptyset$. The result of Dempster's rule indicates that the most credible hypothesis is B, whereas A is not possible to happen, which is not reasonable. The results produced by Yager's rule and DP rule selected in adaptive rule is full of uncertainty, and we even can't make a clear decision from them because of their ways of distributing the mass of conflicting belief. We can get the specific output that most belief focuses on hypothesis A only if PCR5 is selected in the adaptive rule. As we can see, $m_1(.)$ and $m_2(.)$ strongly support the hypothesis A, whereas $m_3(.)$ and $m_4(.)$ strongly support B. It is not easy to be sure what is the true hypothesis. The adaptive rule tends to preserve the earlier decision, since it assumes that $m_1(.)$ and $m_2(.)$ where totally reliable, and then $m_3(.)$ is considered unreliable and thus discounted. When $m_5(.)$ is available, if $m_5(.)$ strongly supports A as with $m_{5A}(.)$, the combination results of all the adaptive rules commit their highest belief in A. If $m_5(.)$ strongly supports B as with $m_{5B}(.)$, the combination results will change and assign the highest belief in B. The results produced by the adaptive rule with selecting combination rules between Dempster and PCR5 are always most specific, which is very useful and helpful for decision-making in real-time. The good performance of adaptive rules lies in the method of evaluating the reliability of sources and the way for automatically selecting suitable combination rules.

VII. Conclusions

An approach for adaptive combination of unreliable sources of evidence has been proposed in this paper for combining sequentially the sources without prior knowledge on their reliabilities. The reliability of each source is evaluated according to its similarity with respect to the others which is measured by an average distance of similarity. When a source is not reliable enough, its bba is discounted to diminish its influence in the fusion process and on decision-making. Before the fusion of the sources, the suitable combination rule is selected depending on the mass of conflicting belief $m_\oplus(\emptyset)$ and the compromise between the computational burden and the specificity of the result one wants to deal with. Whenever $m_\oplus(\emptyset)$ is below the tolerance threshold, Dempster's rule can be chosen as a good rule of combination for such a compromise. Otherwise, Yager's rule, DP rule, or PCR5 must be selected to avoid to get counterintuitive results. The choice among these three rules depends on the application and the acceptable risk in decision-making errors. PCR5 rule is very appropriate to use in general for decision-making because it provides the most specific fusion results, but it requires more computational resources than other rules. If we want to keep uncertain results and don't necessarily need a very

specific decision in case of high conflict between sources, Yager's rule or DP rule can be selected instead. Our numerical example shows the interest of the proposed approach. The main difficulty however lies in the tuning of the thresholds ω_d, $\omega_\emptyset$ and the parameter n involved in its implementation. These parameters must be selected by experience depending on the application. This approach was based on Shafer's model, but could be extended to other models proposed in DSmT.

Acknowledgement

This work was supported by a grant from the National Natural Science Foundation of China(NO.60634030,60775012).

REFERENCES

[1] Y. Deng, W.K. Shi, Z.F. Zhu, Q. Liu, *Combining belief functions based on distance of evidence*, Decision Support Systems, Vol. 38, No. 3, pp. 489–493, 2004.

[2] D. Dubois, H. Prade, *On the unicity of Dempster's rule of combination*, International Journal of Intelligent Systems, Vol. 1, No. 2, pp. 133–142, 1986.

[3] A.L. Jousselme, D. Grenier, E. Bossé, *A new distance between two bodies of evidence*, Information Fusion, Vol. 2, No. 1, pp. 91–101, 2001.

[4] W. Liu, *Analyzing the degree of conflict among belief functions*, Artificial Intelligence, Vol. 170, No. 11, pp. 909–924, 2006.

[5] A. Martin, A.L. Jousselme, C. Osswald, *Conflict measure for the discounting operation on belief functions*, in Proc. of Fusion 2008 Conference, Cologne, Germany, July 2008.

[6] C.K. Murphy, *Combining belief functions when evidence conflicts*, Decision Support Systems, Vol. 29, No. 1, pp. 1–9, 2000.

[7] G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, 1976.

[8] F. Smarandache, J. Dezert (Editors), *Advances and Applications of DSmT for Information Fusion*, American Research Press, Rehoboth, Vol.1-3, 2004-2009.

[9] R.R. Yager, *Hedging in the combination of evidence*, Journal of Information & Optimization Sciences, Vol. 4, No. 1, pp. 73–81, 1983.

[10] L.A. Zadeh, *Review of books: A mathematical theory of evidence*, AI Magazine, Vol. 5, No. 3, pp. 81–83, 1984.

TABLE I

COMBINATION RESULTS BY DIFFERENT RULES

	$\mathbf{m}_1^2$	$\mathbf{m}_1^3$	$\mathbf{m}_1^4$	$\mathbf{m}_1^{5A}$	$\mathbf{m}_1^{5B}$
Dempster's rule	$m(A)=0.7826$ $m(B)=0.1413$ $m(C)=0.0435$ $m(B\cup C)=0.0326$	$m(B)=0.9536$ $m(C)=0.0464$	$m(B)=0.9536$ $m(C)=0.0464$	$m(B)=0.9536$ $m(C)=0.0464$	$m(B)=0.9867$ $m(C)=0.0133$
Yager's rule	$m(A)=0.36$ $m(B)=0.065$ $m(C)=0.02$ $m(B\cup C)=0.015$ $m(\Theta)=0.54$	$m(B)=0.558$ $m(C)=0.0575$ $m(\Theta)=0.3845$	$m(B)=0.4799$ $m(C)=0.2046$ $m(B\cup C)=0.0385$ $m(\Theta)=0.2770$	$m(A)=0.1385$ $m(B)=0.1296$ $m(A\cup B)=0.0692$ $m(C)=0.0885$ $m(A\cup C)=0.0415$ $m(\Theta)=0.5327$	$m(B)=0.5993$ $m(A\cup B)=0.0692$ $m(C)=0.0827$ $m(B\cup C)=0.0473$ $m(\Theta)=0.2015$
DP rule	$m(A)=0.36$ $m(B)=0.065$ $m(A\cup B)=0.24$ $m(C)=0.02$ $m(A\cup C)=0.16$ $m(B\cup C)=0.035$ $m(\Theta)=0.12$	$m(B)=0.414$ $m(A\cup B)=0.324$ $m(C)=0.0335$ $m(A\cup C)=0.036$ $m(B\cup C)=0.0245$ $m(\Theta)=0.168$	$m(B)=0.4925$ $m(C)=0.1249$ $m(B\cup C)=0.2206$ $m(\Theta)=0.1620$	$m(A)=0.081$ $m(B)=0.1783$ $m(A\cup B)=0.2868$ $m(C)=0.1026$ $m(A\cup C)=0.0867$ $m(B\cup C)=0.0493$ $m(\Theta)=0.2153$	$m(B)=0.6897$ $m(A\cup B)=0.0405$ $m(C)=0.0695$ $m(B\cup C)=0.1691$ $m(\Theta)=0.0312$
PCR5 rule	$m(A)=0.7734$ $m(B)=0.1273$ $m(C)=0.0653$ $m(B\cup C)=0.0340$	$m(A)=0.3902$ $m(B)=0.5814$ $m(C)=0.0284$	$m(A)=0.1942$ $m(B)=0.5733$ $m(C)=0.2245$ $m(B\cup C)=0.008$	$m(A)=0.4025$ $m(B)=0.4154$ $m(A\cup B)=0.0296$ $m(C)=0.1346$ $m(A\cup C)=0.0178$ $m(B\cup C)=0.0001$	$m(A)=0.1049$ $m(B)=0.7182$ $m(A\cup B)=0.0296$ $m(C)=0.1334$ $m(B\cup C)=0.0139$
$d_n^{j-1}(\mathbf{m}_j)(n=5)$	0.3813	0.6601	0.4994	0.4115	0.4137
Dempster's rule with discounting	$m(A)=0.7826$ $m(B)=0.1413$ $m(C)=0.0435$ $m(B\cup C)=0.0326$	$m(A)=0.7216$ $m(B)=0.2046$ $m(C)=0.0437$ $m(B\cup C)=0.0301$	$m(B)=0.7565$ $m(C)=0.2255$ $m(B\cup C)=0.018$	$m(B)=0.7608$ $m(C)=0.2392$	$m(B)=0.9194$ $m(C)=0.0770$ $m(B\cup C)=0.0036$
Adaptive rule (Dempster&Yager)	$m(A)=0.7826$ $m(B)=0.1413$ $m(C)=0.0435$ $m(B\cup C)=0.0326$	$m(A)=0.7216$ $m(B)=0.2046$ $m(C)=0.0437$ $m(B\cup C)=0.0301$	$m(B)=0.1261$ $m(C)=0.0376$ $m(B\cup C)=0.0030$ $m(\Theta)=0.8334$	$m(A)=0.4758$ $m(B)=0.0368$ $m(A\cup B)=0.2379$ $m(C)=0.1067$ $m(A\cup C)=0.1427$	$m(B)=0.5550$ $m(A\cup B)=0.2172$ $m(C)=0.0970$ $m(B\cup C)=0.1308$
Adaptive rule (Dempster&DP)	$m(A)=0.7826$ $m(B)=0.1413$ $m(C)=0.0435$ $m(B\cup C)=0.0326$	$m(A)=0.7216$ $m(B)=0.2046$ $m(C)=0.0437$ $m(B\cup C)=0.0301$	$m(B)=0.1261$ $m(A\cup B)=0.3247$ $m(C)=0.0376$ $m(A\cup C)=0.3247$ $m(B\cup C)=0.1147$ $m(\Theta)=0.0722$	$m(A)=0.6232$ $m(B)=0.0765$ $m(A\cup B)=0.126$ $m(C)=0.0988$ $m(A\cup C)=0.0755$	$m(A)=0.1062$ $m(B)=0.5844$ $m(A\cup B)=0.1298$ $m(C)=0.1429$ $m(B\cup C)=0.0367$
Adaptive rule (Dempster&PCR5)	$m(A)=0.7826$ $m(B)=0.1413$ $m(C)=0.0435$ $m(B\cup C)=0.0326$	$m(A)=0.7216$ $m(B)=0.2046$ $m(C)=0.0437$ $m(B\cup C)=0.0301$	$m(A)=0.4634$ $m(B)=0.2975$ $m(C)=0.2273$ $m(B\cup C)=0.0118$	$m(A)=0.7526$ $m(B)=0.1395$ $m(C)=0.1079$	$m(A)=0.2562$ $m(B)=0.6116$ $m(C)=0.1283$ $m(B\cup C)=0.0039$

Fusion of Sources of Evidence with Different Importances and Reliabilities

F. Smarandache
J. Dezert
J.-M. Tacnet

Originally published as Smarandache F., Dezert J., Tacnet J.-M., *Fusion of sources of evidence with different importances and reliabilities*, Fusion 2010, Edinburgh, Scotland, UK, 26-29 July 2010, and reprinted with permission.

Abstract – *This paper presents a new approach for combining sources of evidences with different importances and reliabilities. Usually, the combination of sources of evidences with different reliabilities is done by the classical Shafer's discounting approach. Therefore, to consider unequal importances of sources, if any, a similar reliability discounting process is generally used, making no difference between the notion of importance and reliability. In fact, in multicriteria decision context, these notions should be clearly distinguished. This paper shows how this can be done and we provide simple examples to show the differences between both solutions for managing importances and reliabilities of sources. We also discuss the possibility for mixing them in a global fusion process.*

Keywords: Information fusion, DSmT, discounting, importance, reliability, AHP.

1 Introduction

In many real-life fusion problems, one has to deal with different sources of information arising from human reports, artificial intelligence experts systems and/or physical sensors. The information are usually imprecise, uncertain, incomplete, qualitative or quantitative and possibly conflicting. The task of information fusion is to combine all the information in such a way that one has a better understanding and assessment of the situation of the complex problem under consideration for decision-making support. Several theoretical frameworks have been proposed in the literature (Probability theory, Possibilities theory, Imprecise PT, etc) but the most appealing ones are the theories of belief functions, originally known as Dempster-Shafer Theory (DST) [8] and then extended and refined in Dezert-Smarandache Theory (DSmT) [9] for dealing with qualitative information, for fusioning highly conflicting sources of evidences, for conditioning evidences, etc. Aside the choice of the "best" rule of combination of sources of evidences characterized by their belief functions, more specifically by their basic belief assignments (bba's), or belief masses, the very important problem concerns the possibility that sources involved in the fusion process may not have the same reliability, neither the same importance. The reliability can be seen as an objective property of a source of evidence, whereas the importance of a source is a subjective property of a source expressed by the fusion system designer.

The reliability of a source represents its ability to provide the correct assessment/solution of the given problem. The importance of a source represents somehow the weight of importance granted to the source by the fusion system designer. The reliability and importance represent two distinct notions and the fusion process must be able to deal with these notions. We show in this paper how this can be done efficiently through two discounting techniques using Proportional Conflict Redistribution rules no 5 or no 6 (PCR5 or PCR6) developed in DSmT framework. We will show also that such solution cannot be used in DST framework using Dempster's rule of combination because Dempster's rule doesn't respond to our new importance discounting (it only responds to reliability discounting[1]).

The importance of a source is particularly crucial since it is involved in multi-criteria decision making (MCDM) problems, like in the Analytic Hierarchy Process (AHP) developed by Saaty [6, 7]. That's why it is fundamental to show how the importance can be efficiently managed in evidential reasoning approaches, in particular in DSmT. The fusion system designer is still free to make no differences between importance and reliability and use the classical discounting technique. In general however, one should consider the importance and the reliability as two distinct notions and thus they have to be processed in different ways. This is the purpose of this paper. The application of this technique in DSmT-AHP is presented in [2] and an application of both DSmT and AHP for risk expertise and prevention in mountains has been introduced by Tacnet in [11, 12]

[1] Known as the classical Shafer's discounting, see [8].

and works are still under progress in this field.

This paper is organized as follows. After a brief reminder of basics of DSmT for information fusion and its main fusion rule in section 2, we present the classical discounting technique for combining sources with different reliabilities in section 3. In section 4, we introduce in a new solution for taking into account the possible different importances of sources in the fusion process. Section 5 provides simple examples to show and compare the results obtained from the two discounting approaches. In section 6 we discuss the more general problem where one needs to deal with both reliability and importance at the same level in the fusion process. Conclusions and perspectives are given in section 7.

2 Basics of DSmT

Let $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ be a finite set of n elements θ_i, $i = 1, \ldots, n$ assumed to be exhaustive. Θ corresponds to the frame of discernment of the problem under consideration. In general (unless introducing some integrity constraints), we assume that elements of Θ are non exclusive in order to deal with vague/fuzzy and relative concepts [9], Vol. 2. This is the so-called free-DSm model. In DSmT framework, there is no need in general to work on a refined frame consisting in a discrete finite set of exclusive and exhaustive hypotheses[2] because DSm rules of combination work for any models of the frame, i.e. the free DSm model (no exclusive constraint between θ_i, Shafer's model (all θ_i are exclusive) or any hybrid model (only some θ_i are truly exclusive). The power set 2^Θ is defined as the set of all propositions built from elements of Θ with $\cup$ [8]; Θ generates 2^Θ under $\cup$. The hyper-power set (Dedekind's lattice) D^Θ is defined as the set of all propositions built from elements of Θ with $\cup$ and $\cap$; Θ generates D^Θ under $\cup$ and $\cap$, see [9] Vol. 1 for many detailed examples. The super-power set (Boolean algebra) S^Θ is defined as the set of all propositions built from elements of Θ with $\cup$ and $\cap$ and complement $c(.)$; Θ generates S^Θ under $\cup$, $\cap$ and $c(.)$, see [9] Vol. 3. S^Θ can be seen as the minimal refined frame of Θ. For notation convenience, we use the generic notation G^Θ to represent the fusion space under consideration depending on the application and the underlying model chosen for the frame Θ; which can be either G^Θ can be either 2^Θ, D^Θ or S^Θ. In DST framework, $G^\Theta = 2^\Theta$, whereas in DSmT we usually work with $G^\Theta = D^\Theta$.

A (quantitative) basic belief assignment (bba) expressing the belief committed to the elements of G^Θ by a given source/body of evidence is a mapping function $m(\cdot)$: $G^\Theta \to [0,1]$ such that: $m(\emptyset) = 0$ and $\sum_{A \in G^\Theta} m(A) = 1$. Elements $A \in G^\Theta$ having $m(A) > 0$ are called *focal elements* of the bba $m(.)$. The general belief and plausibility functions are defined respectively in almost the same manner as Shafer in [8], i.e.

$$Bel(A) = \sum_{B \in G^\Theta, B \subseteq A} m(B) \quad (1)$$

$$Pl(A) = \sum_{B \in G^\Theta, B \cap A \neq \emptyset} m(B) \quad (2)$$

In DSmT, the Proportional Conflict Redistribution Rule no. 5 (PCR5) has been proposed as a serious alternative of Dempster's rule [8] for dealing with conflicting belief functions. It has been also clearly shown in [9], Vol. 3, chap. 1 that Smets' rule[3] is not so efficient, nor cogent because it doesn't respond to new information in a global or in a sequential fusion process. Indeed, very quickly Smets' fusion result commits the full mass of belief to the empty set!!! Therefore in applications, some ad-hoc numerical techniques must be used to circumvent this serious drawback. Such problem doesn't occur with PCR5 rule. By construction, other well-known rules like Dubois & Prade, or Yager's rule, and contrariwise to PCR5, increase the non-specificity of the result. An introduction to DSmT and PCR5 fusion rule with justification and several examples can be found in [9], Vol. 3, Chap. 1, freely downloadable from the web.

Definition of PCR5 (for two sources): Let's $m_1(.)$ and $m_2(.)$ be two independent[4] bba's, then the PCR5 rule of combination for two sources of evidence is defined as follows (see [9], Vol. 2 for details, justification and examples): $m_{PCR5}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(A) = \sum_{\substack{X_1, X_2 \in G^\Theta \\ X_1 \cap X_2 = A}} m_1(X_1) m_2(X_2) + \sum_{\substack{X \in G^\Theta \\ X \cap A = \emptyset}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A) + m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A) + m_1(X)} \right] \quad (3)$$

All fractions in (3) having zero denominators are discarded. In DSmT, we consider all propositions/sets in a canonical form. We take the disjunctive normal form, which is a disjunction of conjunctions, and it is unique in Boolean algebra and simplest. For example, $X = A \cap B \cap (A \cup B \cup C)$ it is not in a canonical form, but we simplify the formula and $X = A \cap B$ is in a canonical form. Like most of fusion rules[5], PCR5 is not associative and the *optimal* fusion result is obtained by combining the sources altogether at the same time when possible. Some of PCR5 properties can be found in [1] and it allows non-Bayesian reasoning. An extension of PCR5 for combining qualitative bba's can

[2] Referred as Shafer's model in the literature.

[3] i.e. the non normalized Dempster's rule.

[4] i.e. each source provides its bba independently of the other sources.

[5] Except Dempster's rule, and conjunctive rule in free DSm model.

be found in [9], Vol. 2 & 3.

Basically, the idea of PCR5 is to transfer the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved through this fusion process. For example: consider two bba's $m_1(.)$ and $m_2(.)$, $A \cap B = \emptyset$ for the model of Θ, and $m_1(A) = 0.6$ and $m_2(B) = 0.3$. With PCR5 the partial conflicting mass $m_1(A)m_2(B) = 0.6 \cdot 0.3 = 0.18$ is redistributed to A and B only with respect to the following proportions respectively: $x_A = 0.12$ and $x_B = 0.06$ because the proportionalization requires

$$\frac{x_A}{m_1(A)} = \frac{x_B}{m_2(B)} = \frac{m_1(A)m_2(B)}{m_1(A) + m_2(B)} = \frac{0.18}{0.9} = 0.2$$

Variant of PCR5 (PCR6): The extension and a variant of (3), called PCR6 has been proposed by Martin and Osswald in [9], Vol. 2, for combining $s > 2$ sources and for working in other fusion spaces is presented in [9]. For two sources, PCR6 coincides with PCR5. The difference between PCR5 and PCR6 lies in the way the proportional conflict redistribution is done as soon as three or more sources are involved in the fusion. For example, let's consider three sources with bba's $m_1(.)$, $m_2(.)$ and $m_3(.)$, $A \cap B = \emptyset$ for the model of the frame Θ, and $m_1(A) = 0.6$, $m_2(B) = 0.3$, $m_3(B) = 0.1$. With PCR5 the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR5} = 0.01714$ and $x_B^{PCR5} = 0.00086$ because the proportionalization requires

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B)m_3(B)}$$

that is

$$\frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{0.03} = \frac{0.018}{0.6 + 0.03} \approx 0.02857$$

thus

$$\begin{cases} x_A^{PCR5} = 0.60 \cdot 0.02857 \approx 0.01714 \\ x_B^{PCR5} = 0.03 \cdot 0.02857 \approx 0.00086 \end{cases}$$

With the PCR6 fusion rule, the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR6} = 0.0108$ and $x_B^{PCR6} = 0.0072$ because the PCR6 proportionalization is done as follows:

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_{B,2}^{PCR6}}{m_2(B)} = \frac{x_{B,3}^{PCR6}}{m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B) + m_3(B)}$$

that is

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_{B,2}^{PCR6}}{0.3} = \frac{x_{B,3}^{PCR6}}{0.1} = \frac{0.018}{0.6 + 0.3 + 0.1} = 0.018$$

thus

$$\begin{cases} x_A^{PCR6} = 0.6 \cdot 0.018 = 0.0108 \\ x_{B,2}^{PCR6} = 0.3 \cdot 0.018 = 0.0054 \\ x_{B,3}^{PCR6} = 0.1 \cdot 0.018 = 0.0018 \end{cases}$$

and therefore with PCR6, one gets finally the following redistributions to A and B:

$$\begin{cases} x_A^{PCR6} = 0.0108 \\ x_B^{PCR6} = x_{B,2}^{PCR6} + x_{B,3}^{PCR6} = 0.0054 + 0.0018 = 0.0072 \end{cases}$$

From the implementation point of view, PCR6 is much more simple to implement than PCR5 (see Appendix).

3 Reliability discounting

Reliability refers to information quality while importance refers to subjective preferences of the fusion system designer. The reliability of a source represents its ability to provide the correct assessment/solution of the given problem. It is characterized by a discounting reliability factor, usually denoted α in $[0, 1]$, which should be estimated from statistics when available, or by other techniques [3]. This reliability factor can be context-dependent. For example, if one knows that some sensors do not perform well under bad weather conditions, etc, one will decrease the reliability factor of information arising from that source accordingly. By convention, we usually take $\alpha = 1$ when the source is fully reliable and $\alpha = 0$ if the source is totally unreliable. Reliability of a source is generally considered[6] through Shafer's discounting method [8], p. 252, which consists in multiplying the masses of focal elements by the reliability factor α, and transferring all the remaining discounted mass to the full ignorance Θ. When $\alpha < 1$, such very simple reliability discounting technique discounts all focal elements with the same factor α and it increases the non specificity of the discounted sources since the mass committed to the full ignorance always increases. Mathematically, Shafer's discounting technique for taking into account the reliability factor $\alpha \in [0, 1]$ of a given source with a bba $m(.)$ and a frame Θ is defined by:

$$\begin{cases} m_\alpha(X) = \alpha \cdot m(X), \quad \text{for } X \neq \Theta \\ m_\alpha(\Theta) = \alpha \cdot m(\Theta) + (1 - \alpha) \end{cases} \tag{4}$$

4 Importance discounting

The importance of a source is not the same as its reliability and it can be characterized by an importance factor, denoted β in $[0, 1]$. β factor represents somehow the weight of importance granted to the source by the fusion system designer. The choice of β is usually not related with the reliability of the source and can be chosen to any value in $[0, 1]$ by the designer for his/her

[6] More sophisticated methods have been also proposed, see [4, 5] for example.

own reason. By convention, the fusion system designer will take $\beta = 1$ when he/she wants to grant the maximal importance of the source in the fusion process, and will take $\beta = 0$ if no importance at all is granted to this source in the fusion process. Typically, if one has a pool of experts around a table to take important decision, say politicians, scientific researchers, military officers, etc, it is possible that one wants to grant more importance to the voice of a given politician (say the President) rather than to a military officers or a scientific researcher, even if the scientific researcher is more reliable in his expertise field than other people. Such situations occur frequently in real-life problems. The fusion designer must be able to deal with importance factors in a different way than with reliability factors since they correspond to distinct properties associated with a source of information.

The main question we are concerned in this paper is how to deal with different importances of sources in the fusion process in such a way that a clear distinction is made/preserved between reliability and importance ?

Our preliminary investigations were based on the self/auto-combination of the sources. For example, if one has the importances factors $\beta_1 = 0.7$ for the source s_1 and $\beta_2 = 0.3$ for the source s_2, one could imagine to combine 7 times the bba $m_1(.)$ with itself, combine 3 times the bba $m_2(.)$ with itself, and then combine the resulting auto-fusioned bba's because such combination would reflect somehow the relative importance of the source in the fusion process since $\beta_1/\beta_2 = 0.7/0.3 = 7/3$. Actually such approach is very disputable and cannot be used satisfactorily in practice whatever the fusion rule is adopted. It can be easily shown that the auto-conflict tends quickly to 1 after several auto-fusions [3]. In other words, the combination result of $N \times \beta_1$ bba's $m_1(.)$ with $M \times \beta_2$ bba's $m_2(.)$ is almost the same for any N and M sufficiently large, so that the different importances of sources are not well preserved in such approach. The numerical complexity of such method must be pointed out since it would require to compute possibly many auto-fusions of each source which is a very time-consuming computational task. For example, if $\beta_1 = 0.791$ and $\beta_2 = 0.209$, it would require to combine at least 791 auto-fusions of $m_1(.)$ with 209 auto-fusions of $m_2(.)$!!!

In this paper, we propose a better solution to consider importances of sources. Our new approach can be considered as the dual of Shafer's discounting approach for reliabilities of sources. The idea was originally introduced briefly by Tacnet in [9], Vol.3, Chap. 23, p. 613. It consists to define the importance discounting with respect to the empty set rather than the total ignorance Θ (as done by Shafer in reliability discounting presented in section 3). Such new discounting technique allows to deal easily with sources of different importances and is also very simple to use as it will be shown.

Definition (importance discounting): We define the importance discounting of a source having the importance factor β and asociated bba $m(.)$ by

$$\begin{cases} m_\beta(X) = \beta \cdot m(X), & \text{for } X \neq \emptyset \\ m_\beta(\emptyset) = \beta \cdot m(\emptyset) + (1-\beta) \end{cases} \tag{5}$$

Note that with this importance discounting approach, we allow to deal with non-normal bba since $m_\beta(\emptyset) \geq 0$. The interest of this new discounting is to preserve the specificity of the primary information since all focal elements are discounted with same importance factor and no mass is committed back to partial or total ignorances. Working with positive mass of belief on the empty set is not new and has been introduced in nineties by Smets in his transferable belief model [10]. Here we use the positive mass of the empty set as an intermediate/preliminary step of the fusion process. Clearly when $\beta = 1$ is chosen by the fusion designer, it will mean that the source must take its full importance in the fusion process and so the original bba $m(.)$ is kept unchanged. If the fusion designer takes $\beta = 0$, one will deal with $m_\beta(\emptyset) = 1$ which must be interpreted as a fully non important source. $m(\emptyset) > 0$ is not interpreted as the mass committed to some conflicting information (classical interpretation), nor as the mass committed to unknown elements when working with the open-world assumption (Smets interpretation), but only as the mass of the discounted importance of a source in this particular context.

Before going further, it is worth to note that Dempster's rule cannot deal properly with importance discounted bba's proposed in (5) because our importance discounting technique preserves the specificity of the primary information and Dempster's rule does not make a difference in results when combining $m_1(.)$ with $m_2(.)$ or when combining $m_{\beta_1 \neq 1}(.)$ with $m_{\beta_2 \neq 1}(.)$ due to the way of processing the total conflicting mass of belief. This can be stated as the following theorem:

Theorem 1: Dempster's rule is not responding to the discounting of sources towards the emptyset.
Proof: Let $m_1(.)$ and $m_2(.)$ be two bba's defined on the fusion space $G^\Theta = \{X_1, X_2, \ldots, X_n\}$. Let $m_1(X_i) = a_i$ for all i, with $\sum_{i=1}^n a_i = 1$, and all a_i in $[0,1]$, and let $m_2(X_i) = b_i$ for all i, with $\sum_{i=1}^n b_i = 1$, and all b_i in $[0,1]$. $m_1(\emptyset) = m_2(\emptyset) = 0$. After discounting both $m_1(.)$ and $m_2(.)$ towards the emptyset with β_1 and respectively β_2 in $[0,1]$, we get: $(\beta_1)m_1(X_i) = (\beta_1)a_i$ for all i, with $\sum_{i=1}^n a_i = 1$, and all a_i in $[0,1]$, also $(\beta_1)m_1(\emptyset) = 1 - \beta_1$, and $(\beta_2)m_2(X_i) = (\beta_2)b_i$ for all i, with $\sum_{i=1}^n b_i = 1$, and all b_i in $[0,1]$, also $(\beta_2)m_1(\emptyset) = 1 - \beta_2$. If we apply the conjunctive rule to $m_1(.)$ and $m_2(.)$ we get: $m_{12}(X_i) = c_i$, with $\sum_{i=1}^n c_i = 1$ and c_i in $[0,1]$, where some X_i could be empty intersections. Suppose the

non-empty resulted sets after applying the conjunctive rule are: $X_{i_1},\ldots,\ X_{i_p}$. Then Dempster's rule gives $m_{DS}(X_{i_k}) = m_{12}(X_{i_k})/(m_{12}(X_{i_1})+\ldots+m_{12}(X_{i_p}))$, for $1 \leq k \leq p$. If we apply the conjunctive rule to $(\beta_1)m_1(.)$ and $(\beta_2)m_2(.)$ we get: $(\beta_1)(\beta_2)m_{12}(X_i) = (\beta_1)(\beta_2)c_i$, with $(\beta_1)(\beta_2)c_i$ in $[0,1]$, where some X_i could be empty sets, and $(\beta_1)(\beta_2)m_{12}(\emptyset^7) = 1-(\beta_1)(\beta_2)$. Now Dempster's rule normalizes the conjunctive result of non empty sets by dividing the mass of each nonempty set by the sum of all non-empty sets. The non-empty resulted sets after applying the conjunctive rule are the same: $X_{i_1},\ \ldots,\ X_{i_p}$. Then: $(\beta_1)(\beta_2)m_{DS}(X_{i_k}) = (\beta_1)(\beta_2)m_{12}(X_{i_k})/((\beta_1)(\beta_2)m_{12}(X_{i_1}) + \ldots + (\beta_1)(\beta_2)m_{12}(X_{i_p})) = m_{12}(X_{i_k})/(m_{12}(X_{i_1}) + \ldots + m_{12}(X_{i_p})) = m_{DS}(X_{i_k})$ since the whole fraction is simplified by $(\beta_1)(\beta_2)$, for $1 \leq k \leq p$. Hence Dempster's rule is not responding to the discounting of sources towards the empty set.

From Theorem 1, one understands why such importance discounting technique has never been proposed and used in DST framework and this explains why the classical Shafer's discounting technique (the reliability discounting) has only been largely adopted so far. By using Dempster's rule, the fusion designer has no other choice but to consider importance and reliability as same notions! As it will be shown, the DSmT framework with PCR5 (or PCR6) rule and the importance discounting technique proposed here provides an interesting and simple solution for the fusion of sources with different importances which makes a clear distinction between importances and reliabilities of sources.

Fusion of importance discounted bba's: Based on this new discounting technique, it is however very simple to adapt PCR5 or PCR6 fusion rules for combining the $s \geq 2$ discounted bba's associated with each source i, $i = 1, 2, \ldots s$. It suffices actually to consider the following extension of PCR5, denoted $\text{PCR5}_\emptyset$ and defined by:

- For two sources ($s = 2$): $\forall A \in G^\Theta$ (A may be the empty set too)

$$m_{PCR5_\emptyset}(A) = \sum_{\substack{X_1,X_2\in G^\Theta \\ X_1\cap X_2 = A}} m_1(X_1)m_2(X_2) + \sum_{\substack{X\in G^\Theta \\ X\cap A=\emptyset}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A)+m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A)+m_1(X)}\right] \quad (6)$$

- For $s \geq 2$ sources: $\forall A \in G^\Theta$ (A may be the empty set too)

$$m_{PCR5_\emptyset}(A) = m_{12\ldots s}(A) + \sum_{\substack{2\leq t\leq s \\ 1\leq r_1,\ldots,r_t\leq s \\ 1\leq r_1<r_2<\ldots<r_{t-1}<(r_t=s)}} \sum_{\substack{X_{j_2},\ldots,X_{j_t}\in G^\Theta \\ \{j_2,\ldots,j_t\}\in\mathcal{P}^{t-1}(\{1,\ldots,n\}) \\ A\cap X_{j_2}\cap\ldots\cap X_{j_s}=\emptyset \\ \{i_1,\ldots,i_s\}\in\mathcal{P}^s(\{1,\ldots,s\})}} \frac{(\prod_{k_1=1}^{r_1} m_{i_{k_1}}(A)^2)\cdot[\prod_{l=2}^{t}(\prod_{k_l=r_{l-1}+1}^{r_l} m_{i_{k_l}}(X_{j_l})]}{(\prod_{k_1=1}^{r_1} m_{i_{k_1}}(A)) + [\sum_{l=2}^{t}(\prod_{k_l=r_{l-1}+1}^{r_l} m_{i_{k_l}}(X_{j_l})]} \quad (7)$$

where i, j, k, r, s and t in (7) are integers. $m_{12\ldots s}(A) \triangleq \sum_{\substack{X_1,X_2,\ldots,X_s\in G^\Theta \\ \cap_{i=1}^s X_i = A}} \prod_{i=1}^s m_i(X_i)$ is the conjunctive consensus on A between s sources and where all denominators are different from zero. If a denominator is zero, that fraction is discarded; $\mathcal{P}^k(\{1,2,\ldots,n\})$ is the set of all subsets of k elements from $\{1,2,\ldots,n\}$ (permutations of n elements taken by k), the order of elements doesn't count.

A similar extension can be done for the PCR6 formula for $s > 2$ sources given in [9], Vol. 2. More precisely for any A in G^Θ (A may be the empty set too) we define:

$$m_{PCR6_\emptyset}(A) = m_{12\ldots s}(A) + \sum_{\substack{X_1,X_2,\ldots,X_{s-1}\in G^\Theta \\ X_i\neq A, i\in\{1,2,\ldots,s-1\} \\ (\cap_{i=1}^{s-1} X_i)\cap A=\emptyset}} \sum_{k=1}^{s-1} \sum_{(i_1,i_2,\ldots,i_s)\in P(1,2,\ldots,s)} [m_{i_1}(A) + m_{i_2}(A) + \ldots + m_{i_k}(A)]\times \frac{\prod_{j=1}^{k} m_{i_j}(A)\prod_{p=k+1}^{s-1} m_{i_p}(X_p)}{\sum_{j=1}^{k} m_{i_j}(A) + \sum_{p=k+1}^{s-1} m_{i_p}(X_p)} \quad (8)$$

where $P(1,2,\ldots,s)$ is the set of all permutations of the elements $\{1,2,\ldots,s\}$. It should be observed that $X_1,\ X_2,\ \ldots,\ X_{s-1}$ may be different from each other, or some of them equal and others different, etc.

As a particular case for $s = 3$ sources, one gets for any A in G^Θ (A may be the empty set too):

$$m_{PCR6_\emptyset}(A) = m_{123}(A) + \sum_{\substack{X,Y\in G^\Theta \\ X\neq A, Y\neq A \\ X\cap Y\cap A=\emptyset}} \sum_{(i_1,i_2,i_3)\in P(1,2,3)} \frac{m_{i_1}(A)^2 m_{i_2}(X) m_{i_3}(Y)}{m_{i_1}(A)+m_{i_2}(X)+m_{i_3}(Y)} + [m_{i_1}(A)+m_{i_2}(A)]\cdot\frac{m_{i_1}(A)m_{i_2}(A)m_{i_3}(X)}{m_{i_1}(A)+m_{i_2}(A)+m_{i_3}(X)} \quad (9)$$

[7] i.e. the absolute empty set, not that resulted from set intersections which are empty.

where $m_{123}(A)$ is the mass of the conjunctive consensus on A and $P(1,2,3)$ is the set of all permutations of the elements $\{1,2,3\}$. It should be observed that X may be different or equal to Y.

The difference between formulas (3) and (6) is that $m_{PCR5}(\emptyset) = 0$ whereas $m_{PCR5_\emptyset}(\emptyset) \geq 0$. Of course, since we usually need to work with normal bba's for decision-making support, the result $m_{PCR5_\emptyset}(.)$, or $m_{PCR6_\emptyset}(.)$, of the fusion of discounted masses $m_{\beta_i}(.)$ will be normalized by redistributing the mass of belief committed to the empty set to the other focal elements and proportionally to their masses (see next example).

5 Example

For convenience and simplicity, and due to space limitation constraint, we give a very simple example working on the classical power set 2^Θ since most of readers familiar belief functions usually work with this fusion space. **Example 1**: Let's consider $\Theta = \{A, B\}$, Shafer's model, and two sources with respectively bba's $m_1(.)$ and $m_2(.)$ given by $m_1(A) = 0.8$, $m_1(B) = 0.2$ and $m_2(A) = 0.4$, $m_2(B) = 0.6$.

- Case 1 (no importance discounting): Let's consider that $\beta_1 = 1$ and $\beta_2 = 1$, i.e. the sources must have the same maximal importance in the fusion rule. In that case, one gets: $m_{\beta_1}(.) = m_1(.)$ and $m_{\beta_2}(.) = m_2(.)$ and the bba's are actually not discounted. The conjunctive rule gives $m_{12}(A) = 0.32$, $m_{12}(B) = 0.12$ and the mass $m_{12}(A \cap B = \emptyset) = 0.56$ is redistributed back to A and B proportionally to their masses following the PCR5 principle explained in section 2. We get the following result:

	$m_{\beta_1=1}(.)$	$m_{\beta_2=1}(.)$	$m_{12}(.)$	$m_{PCR5}(.)$
$\emptyset$	0	0	0.56	0
A	0.8	0.4	0.32	0.64
B	0.2	0.6	0.12	0.36

Table 1: PCR5 fusion of $m_{\beta_1=1}(.)$ with $m_{\beta_2=1}(.)$.

- Case 2 (with importance discounting): Let's take now the importances factors $\beta_1 = 0.2$ and $\beta_2 = 0.8$ (note that in general we don't need to force the sum of β_i to be one, unless one wants to deal with relative importances between sources). Applying importance discounting technique and normalization of $m_{PCR5_\emptyset}$, denoted $m^{norm}_{PCR5_\emptyset}(.)$, one gets:

	$m_{\beta_1=0.2}(.)$	$m_{\beta_2=0.8}(.)$	$m_{12}(.)$	$m^{norm}_{PCR5_\emptyset}(.)$
$\emptyset$	0.80	0.20	0.9296	0
A	0.16	0.32	0.0512	0.43
B	0.04	0.48	0.0192	0.57

Table 2: PCR5 fusion of $m_{\beta_1=0.2}(.)$ with $m_{\beta_2=0.8}(.)$.

Clearly, one sees in Table 2 the strong impact of the importance discounting on the result with respect to what we obtain in Table 1 (i.e. without importance discounting). Note also that the difference is very different to what we would have obtained by taking $\alpha_1 = 0.2$ and $\alpha_2 = 0.8$ and using the reliability discounting approach as seen in Table 3.

	$m_{\alpha_1=0.2}(.)$	$m_{\alpha_2=0.8}(.)$	$m_{12}(.)$	$m_{PCR5}(.)$
$\emptyset$	0	0	0.0896	0
A	0.16	0.32	0.3392	0.3698
B	0.04	0.48	0.4112	0.4702
$A \cup B$	0.80	0.20	0.1600	0.1600

Table 3: PCR5 fusion of $m_{\alpha_1=0.2}(.)$ with $m_{\alpha_2=0.8}(.)$.

By comparing Table 2 with Table 3, one sees the clear difference in results obtained by these two discounting techniques which is normal.

6 Reliability and importance

In this section, we examine the possibility to take into account both the reliabilities α_i and the importances β_i of given sources of evidence characterized by their bba's $m_i(.)$, $i = 1, 2, \ldots, s$. The main question is how to deal with these two distinct discounting factors since in general, but when $\alpha_i = \beta_i = 1$, the reliability and importance discounting approaches do not commute. Indeed, it can be easily verified (see in next example) that $m_{\alpha_i,\beta_i}(.) \neq m_{\beta_i,\alpha_i}(.)$ whenever $\alpha_i \neq 1$ and $\beta_i \neq 1$. $m_{\alpha_i,\beta_i}(.)$ denotes the reliability discounting of $m_i(.)$by α_i followed by the importance discounting of $m_{\alpha_i}(.)$ by β_i which explains the notation α_i, β_i used in index. Similarly, $m_{\beta_i,\alpha_i}(.)$ denotes the importance discounting of $m_i(.)$ by β_i followed by the reliability discounting of $m_{\beta_i}(.)$ by α_i. To deal both with reliabilities and importances factors and because of the non commutativity of these discountings, we propose to proceed the fusion of the sources in a three-steps process as follows:

Method 1: Step 1: Apply reliability and then importance discountings to get $m_{\alpha_i,\beta_i}(.)$, $i = 1, \ldots, s$ and combine them with $PCR5_\emptyset$ or $PCR6_\emptyset$ and normalize the resulting bba; Step 2: Apply importance and then reliability discountings to get $m_{\beta_i,\alpha_i}(.)$, $i = 1, \ldots, s$ and combine them with $PCR5_\emptyset$ or $PCR6_\emptyset$ and normalize the resulting bba; Step 3 (mixing/averaging): Combine the resulting bba's of Steps 1 and 2 using the arithmetic mean operator[8].

Method 2: Another simplest method which could be useful for intermediate traceability in some applications would consist to change Steps 1 & 2 by Step 1': Apply reliability discounting only to get $m_{\alpha_i}(.)$ and combine them with PCR5 or PCR6; Step 2': Apply importance discounting only to get $m_{\beta_i}(.)$ and combine them with

[8]Other combination rules could be used also like PCR5 or PCR6, etc., but we don't see solid justification to use them again and they require more computations than the simple arithmetic mean.

$PCR5_\emptyset$ or $PCR6_\emptyset$ and normalize the result; Step 3' same as Step 3. Due to space limitation, only Method 1 is briefly illustrated in the following simple example.

Example 2: Let's take $\Theta = \{A, B, C\}$, Shafer's model, three sources $m_1(.)$, $m_2(.)$ and $m_3(.)$ given in next table and assume that their reliability factors are $\alpha_1 = 0.8$, $\alpha_2 = 0.5$, and $\alpha_3 = 0.2$ and their importance factors are $\beta_1 = 0.9$, $\beta_2 = 0.3$ and $\beta_3 = 0.6$.

	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\emptyset$	0	0	0
A	0.8	0.4	0.1
B	0	0.3	0.3
$A \cup B$	0.1	0.2	0
C	0	0	0.5
$A \cup C$	0.1	0	0
$B \cup C$	0	0.1	0
$A \cup B \cup C$	0	0	0.1

Table 4: Sources of evidences.

By applying reliability followed by importance discounting, and by applying importance followed by reliability discounting, one gets:

	$m_{\alpha_1,\beta_1}(.)$	$m_{\alpha_2,\beta_2}(.)$	$m_{\alpha_3,\beta_3}(.)$
$\emptyset$	0.1000	0.7000	0.4000
A	00.5760	0.0600	0.0120
B	0	0.0450	0.0360
$A \cup B$	0.0720	0.0300	0
C	0	0	0.0600
$A \cup C$	0.0720	0	0
$B \cup C$	0	0.0150	0
$A \cup B \cup C$	0.1800	0.1500	0.4920

Table 5: Reliability-Importance discounting.

	$m_{\beta_1,\alpha_1}(.)$	$m_{\beta_2,\alpha_2}(.)$	$m_{\beta_3,\alpha_3}(.)$
$\emptyset$	0.0800	0.3500	0.0800
A	0.5760	0.0600	0.0120
B	0	0.0450	0.0360
$A \cup B$	0.0720	0.0300	0
C	0	0	0.0600
$A \cup C$	0.0720	0	0
$B \cup C$	0	0.0150	0
$A \cup B \cup C$	0.2000	0.5000	0.8120

Table 6: Importance-Reliability discounting.

The normalized results of the $PCR5_\emptyset$ fusion of $m_{\alpha_i,\beta_i}(.)$ for $i = 1, 2, 3$ (Step 1) and $PCR5_\emptyset$ fusion of $m_{\beta_i,\alpha_i}(.)$ for $i = 1, 2, 3$ (Step 2) is given in next Table 7 with their arithmetic mean $\bar{m}_{PCR5}(.)$ (Step 3).

	Step 1	Step 2	Step 3
	$m^{norm}_{PCR5_{\emptyset,\alpha,\beta}}(.)$	$m^{norm}_{PCR5_{\emptyset,\beta,\alpha}}(.)$	$\bar{m}_{PCR5}(.)$
$\emptyset$	0	0	0
A	0.5741	0.4927	0.5334
B	0.0254	0.0244	0.0249
$A \cup B$	0.0311	0.0464	0.0388
C	0.0182	0.0182	0.0182
$A \cup C$	0.0233	0.0386	0.0310
$B \cup C$	0.0032	0.0032	0.0032
$A \cup B \cup C$	0.3247	0.3765	0.3506

Table 7: Results of Steps 1, 2 & 3.

7 Conclusions

The proposition of two different discounting techniques is an important contribution to consider both preferences and reliability in fusion problems for decision making purposes. In this paper, we have proposed a new solution for taking into account the different importances of sources of evidence in their combination. We have shown the clear distinction between the classical reliability discounting technique and our new importance discounting method which can be used with extensions of PCR5 and PCR6 fusion rules developed in DSmT framework. It has been shown also that Dempster's rule cannot be applied satisfactorily with this importance discounting approach contrariwise to PCR5 and PCR6 rules. The importance and reliability can now be distinguished in the fusion of sources which introduces a link with Multi-Criteria Decision Problems in the fusion of sources of information. Applications of these techniques for risk prevention against natural catastrophes in mountains are under progress and results will be published in forthcoming publications.

References

[1] J. Dezert, F. Smarandache, *Non Bayesian conditioning and deconditioning*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[2] J. Dezert, J.-M. Tacnet, M. Batton-Hubert, F. Smarandache, *Multi-criteria decision making based on DSmT-AHP*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[3] A. Martin, A.-L. Jousselme, C. Osswald, *Conflict measure for the discounting operation on belief functions*, Proc. of Fusion 2008 Int. Conference, Cologne, Germany, July 2008.

[4] D. Mercier, B. Quost, T. Denoeux, *Contextual discounting of belief functions*, Proc. of EC-QSARU2005, pp. 552–562, Barcelona, July 2005.

[5] D. Mercier, T. Denoeux, M.-H. Masson, *Refined sensor tuning in the belief function framework using contextual discounting*, Proc. of IPMU2006, Vol II, pp. 1443-1450, Paris, France, July 2006.

[6] T.L. Saaty, *The Analytical Hierarchy Process*, McGraw Hill, New York, 1980.

[7] T.L. Saaty, *Fundamentals of decision making and priority theory with the analytic hierarchy process*, Vol. VI of the AHP series, RWL Publications, Pittsburgh, PA, USA.

[8] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[9] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3, American Research Press, 2004–2009. http://fs.gallup.unm.edu/DSmT.htm

[10] Ph. Smets Ph., *The Combination of Evidence in the Transferable Belief Model*, IEEE Trans. PAMI 12, pp. 447–458, 1990.

[11] J.-M. Tacnet, M. Batton-Hubert, J. Dezert, *Information fusion for natural hazards in mountains*, Chapter 23 of [9], Vol.3, June 2009.

[12] J.-M. Tacnet, M. Batton-Hubert, J. Dezert, *A two-step fusion process for multi-criteria decision applied to natural hazards in mountains*, Int. Workshop on Belief Functions, Brest, April 2010.

Appendix: Matlab^{TM}code listings for PCR5 and PCR6

For convenience, we provide two Matlab^{TM}routines for PCR5 and PCR6 for the fusion of $s \geq 2$ sources for working with 2^{Θ}, i.e. working with Shafer's model. Some adaptations need to be done to work on other fusion spaces and to work with $PCR5_{\emptyset}$ and $PCR6_{\emptyset}$. No verification of input is done in the routines. It is assumed that the input matrix BBA is correct, both in dimension and in content. No attempt for fast computation, nor memory optimization is done in these very simple and basic codes. The derivation of all possible combinations in the loop with `combvec(Combinations,vec)` instruction is a very time-consuming task when the size of the problem increases and should be done once outside the routines. The j-th column of the BBA input matrix corresponds to the (vertical) bba vector $m_j(.)$ associated with the j-th source s_j. Each element of a BBA matrix is in [0,1] and the sum of each column must be one. If N is the cardinality of the frame Θ and if S is the number of sources, then the size of the BBA input matrix is $((2^N) - 1)) \times S$. Each column of the BBA matrix must use the following binary encoding of elements[9] of $2^{\Theta} \setminus \{\emptyset\}$. For example, if $\Theta = \{A, B, C\}$, then binary sequence $001 = A$, $010 = B$, $011 = A \cup B$, $\ldots$, $111 = A \cup B \cup C$. These codes can be used and shared for free for research purposes only. Commercial uses of these codes, or any adaptation of them, is not allowed without written agreement of the author. The use of these codes are at the own risk of the user.

[9] Since one always considers normal input bba's such that $m_j(\emptyset) = 0$, $j = 1, \ldots S$, one doesn't need to store these values in the BBA matrix. For $PCR5_{\emptyset}$ and $PCR6_{\emptyset}$ however, one needs to include as first row of BBA the $m_j(\emptyset) \geq 0$ resulting from importance discounting of the sources and make a proper adaptation of indexes in the routines.

File : PCR5fusion.m

```
function [mPCR5,TotalConflict]=PCR5fusion(BBA)
% Author and copyrights: Jean Dezert
% Input: BBA matrix
% Output: mPCR5 = resulting bba after fusion with PCR5
% TotalConflict = level of total conflict between sources
NbrSources=size(BBA,2);
CardTheta=log2(size(BBA,1)+1);
if(NbrSources==1)
mPCR5=BBA(:,1);TotalConflict=0;return
end
Card2PowerTheta=2^(CardTheta)-1;
% All possible combinations
vec=[1:Card2PowerTheta];
Combinations=vec;
for s=1:NbrSources-1
Combinations=combvec(Combinations,vec);
end
Combinations=Combinations';
mPCR5=zeros(Card2PowerTheta,1);
TotalConflict=0;
NbrComb=size(Combinations,1);
for c=1:NbrComb
PC=Combinations(c,:);
mConj=zeros(1,NbrSources);
for s=1:NbrSources
mConj(s)=BBA(PC(s),s);
end
massConj=prod(mConj,2);
if(massConj>0)
% Check if this is a real partial conflict or not
Intersections=PC(1);
for s=2:NbrSources
X=PC(s);
Intersections=bitand(Intersections,X);
end
if(Intersections~=0) % the intersection is not empty
mPCR5(Intersections)=mPCR5(Intersections)+massConj;
else % the intersection is empty
TotalConflict=TotalConflict+massConj;
% Let's apply PCR5 rule principle
UQ=unique(PC);
Proportions=0*UQ;
DenPCR5=0;
for u=1:size(UQ,2)
SamePropositions=find(PC==UQ(u));
MassProd=prod(mConj(SamePropositions));
Proportions(u)= MassProd*massConj;
DenPCR5=DenPCR5+MassProd;
end
Proportions=Proportions/DenPCR5;
% PCR5 redistribution
for u=1:size(UQ,2)
mPCR5(UQ(u))=mPCR5(UQ(u))+Proportions(u);
end, end, end, end, return
```

File : PCR6fusion.m

```
function [mPCR6,TotalConflict]=PCR6fusion(BBA)
% Author and copyrights: Jean Dezert
% Input: BBA matrix
% Output: mPCR6 = resulting bba after fusion with PCR6
% TotalConflict = level of total conflict between sources
NbrSources=size(BBA,2);
CardTheta=log2(size(BBA,1)+1);
if(NbrSources==1)
mPCR6=BBA(:,1);
TotalConflict=0;
return
end
Card2PowerTheta=2^(CardTheta)-1;
% All possible combinations
vec=[1:Card2PowerTheta];
Combinations=vec;
for s=1:NbrSources-1
Combinations=combvec(Combinations,vec);
end
Combinations=Combinations';
mPCR6=zeros(Card2PowerTheta,1);
TotalConflict=0;
NbrComb=size(Combinations,1);
for c=1:NbrComb
PC=Combinations(c,:);  % particular combination
mConj=zeros(1,NbrSources);
for s=1:NbrSources
mConj(s)=BBA(PC(s),s);
end
massConj=prod(mConj,2);
if(massConj>0)
Intersections=PC(1);
for s=2:NbrSources
X=PC(s);
Intersections=bitand(Intersections,X);
end
if(Intersections~=0) % intersection not empty
mPCR6(Intersections)=mPCR6(Intersections)+massConj;
else % empty intersection
TotalConflict=TotalConflict+massConj;
% PCR6 rule principle
for s=1:NbrSources
Proportion= mConj(s)*(massConj/(sum(mConj,2)));
% Redistribution back to element PC(s)
mPCR6(PC(s))=mPCR6(PC(s))+Proportion;
end, end, end, end, return
```

Is Entropy Enough to Evaluate the Probability Transformation Approach of Belief Function?

Deqiang Han
Jean Dezert
Chongzhao Han
Yi Yang

Originally published as Han D., Dezert J., Han C., *Is Entropy Enough to Evaluate the Probability Transformation Approach of Belief Function?*, Fusion 2010, Edinburgh, Scotland, UK, 26-29 July 2010, and reprinted with permission.

Abstract − *In Dempster-Shafer Theory (DST) of evidencee and transferable belief model (TBM), the probability transformation is necessary and crucial for decision-making. The evaluation of the quality of the probability transformation is usually based on the entropy or the probabilistic information content (PIC) measures, which are questioned in this paper. Another alternative of probability transformation approach is proposed based on the uncertainty minimization to verify the rationality of the entropy or PIC as the evaluation criteria for the probability transformation. According to the experimental results based on the comparisons among different probability transformation approaches, the rationality of using entropy or Probabilistic Information Content (PIC) measures to evaluate probability transformation approaches is analyzed and discussed.*

Keywords: TBM, uncertainty, pignistic probability transformation, evidence theory, decision-making.

1 Introduction

Evidence theory, as known as Dempster-Shafer Theory (DST) [1, 2] can reason with imperfect information including imprecision, uncertainty, incompleteness, etc. It is widely used in many fields in information fusion. There are also some drawbacks and problems in evidence theory, i.e. the high computational complexity, the counter-intuitive behaviors of Dempster's combination rule and the decision-making in evidence theory, etc. Several modified, refined or extended models were proposed to resolve the problems aforementioned, such as transferable belief model (TBM) [3] proposed by Philippe Smets and Dezert-Smarandache Theory (DSmT) [4] proposed by Jean Dezert and Florentin Smarandache, etc.

The goal of uncertainty reasoning is the decision-making. To take a decision, the belief assignment values for a compound focal element should be at first assigned to the singletons. So the probability transformation from belief function is crucial for the decision-making in evidence theory. The research on probability transformation has attracted more attention in recent years.

The most famous probability transformation in evidence theory is the pignistic probability transformation (PPT) in TBM. TBM has two levels including credal level and pignistic level. At the credal level, beliefs are entertained, combined and updated while at the pignistic level, the PPT maps the beliefs defined on subsets to the probability defined on singletons, then a classical probabilistic decision can be made. In PPT, belief assignment values for a compound focal element are equally assigned to the singletons belonging to the focal element. In fact, PPT is designed according to principle of minimal commitment, which is somehow related with uncertainty maximization. But the goal of information fusion at decision-level is to reduce the uncertainty degree. That is to say more uncertainty might not be helpful for the decision. PPT uses equal weights when splitting masses of belief of partial uncertainties and redistributing them back to singletons included in them. Other researchers also proposed some modified probability transformation approaches [5–13] to assign the belief assignment values of compound focal elements to the singletons according to some ratio constructed based on some available information. The typical approaches include the Sudano's probabilities [8] and the Cuzzolin's intersection probability [13], etc. In the framework of DSmT, another probability transformation approach was proposed, which is called DSmP [9]. DSmP takes into account both the values of

the masses and the cardinality of focal elements in the proportional redistribution process. DSmP can also be used in Shafer's model within DST framework.

In almost all the research works on probability transformations, the entropy or Probabilistic Information Content (PIC) criteria are used to evaluate the probability transformation approaches. Definitely for the purpose of decision, less uncertainty should be better to make a more clear and solid decision. But does the probability distribution generated from belief functions with less uncertainty always rational or always be benefit to the decision? We do not think so. In this paper, an alternative probability transformation approach based on the uncertainty minimization is proposed. The objective function is established based on the Shannon entropy and the constraints are established based on the given belief and plausibility functions. The experimental results based on some provided numerical examples show that the probability distributions generated based on the proposed alternative approach have the least uncertainty degree when compared with other approaches. When using the entropy or PIC to evaluate the proposed probability transformation approach, the probability distribution with the least uncertainty seemingly should be the optimal one. But some risky and strange results can be derived in some cases, which are illustrated in some numerical examples. It can be concluded that the entropy or PIC, i.e. the uncertainty degree might not be enough to evaluate the probability transformation approach. In another word, the entropy or PIC might not be used as the only criterion to make the evaluation.

2 Basics of evidence theory and probability transformation

2.1 Basics of evidence theory

In Dempster-Shafer theory [2], the elements in the frame of discernment (FOD) Θ are mutually exclusive. Define the function $m : 2^{\Theta} \rightarrow [0, 1]$ as the basic probability assignment (BPA, also called mass function), which satisfies:

$$\sum\nolimits_{A\subseteq\Theta} m(A) = 1,\ m(\emptyset) = 0 \tag{1}$$

Belief function and plausibility function are defined respectively in (2) and (3):

$$Bel(A) = \sum\nolimits_{B\subseteq A} m(B) \tag{2}$$

$$pl(A) = \sum\nolimits_{A\cap B\neq\emptyset} m(B) \tag{3}$$

and Dempster's rule of combination is defined as follows: $m_1, m_2, ..., m_n$ are n mass functions, the new combined evidence can be derived based on (4)

$$m(A) = \begin{cases} 0, & A = \emptyset \\ \dfrac{\sum\limits_{\cap A_i = A} \prod\limits_{1\leq i\leq n} m_i(A_i)}{\sum\limits_{\cap A_i \neq \emptyset} \prod\limits_{1\leq i\leq n} m_i(A_i)}, & A \neq \emptyset \end{cases} \tag{4}$$

Dempster's rule of combination is used in DST to accomplish the fusion of bodies of evidence. But the final goal of the information fusion at decision-level is to make the decision. The belief function (or BPA, plausibility function) should be transformed to the probability, before the probability-based decision-making. Although there are also some research works on making decision directly based on belief function or BPA [14], probability-based decision methods are the development trends of uncertainty reasoning and theories [15]. This is because the two-level reasoning and decision structure proposed by Smets in his TBM is appealing.

2.2 Pignistic transformation

As a type of probability transformation approach, the classical pignistic probability in TBM framework was coined by Philippe Smets. TBM is a subjective and non probabilistic interpretation of evidence theory. It extends the evidence theory to the open–world propositions and it has a range of tools for handling belief functions including discounting and conditioning, etc. At the credal level of TBM, beliefs are entertained, combined and updated while at the pignistic level, beliefs are used to make decisions by transforming beliefs to probability distribution based on pignistic probability transformation (PPT). The basic idea of the pignistic transformation consists in transferring the positive belief of each compound (or nonspecific) element onto the singletons involved in that element split by the cardinality of the proposition when working with normalized BPAs.

Suppose that $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$ is the FOD. The PPT for the singletons is illustrated as follows [3]:

$$\mathrm{BetP}_m(\theta_i) = \sum_{\theta_i\in B,\ B\subseteq 2^{\Theta}} \frac{m(B)}{|B|} \tag{5}$$

where 2^{Θ} is the power set of the FOD. Based on the pignistic probability derived, the corresponding decision can be made.

But in fact, PPT is designed according to the idea being similar to uncertainty maximization. In general, the PPT is just a simple averaging operation. The mass value is not assigned discriminately to the different singletons involved. But for information fusion, the aim is to reduce the degree of uncertainty and to gain a more consolidated and reliable decision result. The high uncertainty in PPT might be not helpful for the decision. Several researchers aim to modify the traditional PPT. Some typical modified probability transformation approaches are as follows.

1) Sudano's probabilities: Sudano [8] proposed some interesting alternatives to PPT denoted by PrPl, PrNPl, PraPl, PrBel and PrHyb, respectively. Sudano uses different kinds of mappings either proportional to the plausibility, to the normalized plausibility, to all plausibilities and to the belief, respectively or a hybrid mapping.

2) Cuzzolin's intersection probability: In the framework of DST, Fabio Cuzzolin [13] proposed another type of transformation. From a geometric interpretation of Dempster's combination rule, an intersection probability measure was proposed from the proportional repartition of the total non specific mass (TNSM) by each contribution of the non-specific masses involved in it.

3) DSmP: Dezert and Smarandache proposed the DSmP as follows: Suppose that the FOD is $\Theta = \{\theta_1, ..., \theta_n\}$, the $\mathrm{DSmP}_\varepsilon(\theta_i)$can be directly obtained by:

$$\mathrm{DSmP}_\varepsilon(\theta_i) = m(\{\theta_i\}) + (m(\{\theta_i\}) + \varepsilon)\cdot \left(\sum_{\substack{X\in 2^\Theta \\ \theta_i \subset X \\ |X|\geq 2}} \frac{m(X)}{\sum\limits_{\substack{Y\in 2^\Theta \\ Y\subset X \\ |Y|=1}} m(Y)+\varepsilon\cdot|X|} \right) \quad (6)$$

In DSmP, both the values of the mass assignment and the cardinality of focal elements are used in the proportional redistribution process. DSmP does an *improvement* of all Sudano, Cuzzolin, and BetP formulas, in the sense that DSmP mathematically makes a more accurate redistribution of the ignorance masses to the singletons involved in ignorance. DSmP works in both theories: DST and DSmT as well.

There are still some other definitions on modified PPT such as the iterative and self-consistent approach PrScP proposed by Sudano in [5] and a modified PrScP in [12]. Although the approaches aforementioned are different, all the probability transformation approaches are evaluated based on the degree of uncertainty. Less uncertainty means that the corresponding probability transformation result is better. According to such a idea, the probability transformation approach should attempt to enlarge the belief differences among all the propositions and thus to derive a more reliable decision result. Is this definitely rational? Is the uncertainty degree always proper or enough to evaluate the probability transformation? In the following section, some uncertainty measures are analyzed and an alternative probability transformation approach based on uncertainty minimization is proposed to verify the rationality of the uncertainty degree as the criteria for evaluating the probability transformation.

3 An alternative probability transformation based on uncertainty minimization

3.1 Evaluation criteria for probability transformation

The metrics depicting the strength of a critical decision by a specific probability distribution are introduced as follows:

1) Normalized Shannon entropy

Suppose that p_θ is a probability distribution, where $\theta \in \Theta$, $|\Theta| = N$ and the $|\Theta|$ represents the cardinality of the FOD Θ. The evaluation criterion for the probability distribution derived based on different probability transformation is as follows [12].

$$\mathrm{E_H} = \frac{-\sum\limits_{\theta\in\Theta} p_\theta \log_2(p_\theta)}{\log_2 N} \quad (7)$$

The dividend in (7) is the Shannon entropy and the divisor in (7) is maximum value of the Shannon entropy for $\{p_\theta|\theta \in \Theta\}$,$|\Theta| = N$. Obviously E_H is normalized. The larger the E_H is, the larger the degree of uncertainty is. The less the E_H is, the less the degree of uncertainty is. When E_H= 0, there is only one hypothesis has a probability value of 1 and the rest has 0, the agent or system can make decision correctly. When E_H= 1, it is impossible to make a correct decision, because all the $p_\theta, \forall\theta \in \Theta$ are equal.

2) Probabilistic Information Content

Probabilistic Information Content (PIC) criterion is an essential measure in any threshold-driven automated decision system. A PIC value of one indicates the total knowledge to make a correct decision.

$$\mathrm{PIC}(P) = 1 + \frac{1}{\log_2 N} \cdot \sum_{\theta\in\Theta} p_\theta \log_2(p_\theta) \quad (8)$$

Obviously, $\mathrm{PIC} = 1 - \mathrm{E_H}$. The PIC is the dual of the normalized Shannon entropy. A PIC value of zero indicates that the knowledge to make a correct decision does not exist (all the hypotheses have an equal probability value), i.e. one has the maximal entropy.

As referred above, for information fusion at decision-level, the uncertainty seemingly should be reduced as much as possible. The less the uncertainty in probability measure is, the more consolidated and reliable decision can be made. Suppose such a viewpoint is always right and according to such an idea, an alternative probability transformation of belief function is proposed.

3.2 Probability transformation of belief function based on uncertainty minimization

To accomplish the probability transformation, the belief function (or the BPA, the plausibility function)

should be available. The relationship between the probability and the belief function are analyzed as follows.

Based on the viewpoint of Dempster and Shafer, the belief function can be considered as a lower probability and the plausibility can be considered as an upper probability. Suppose that $p_\theta \in [0,1]$ is a probability distribution, where $\theta \in \Theta$. For a belief function defined on FOD Θ, suppose that $B \in 2^\Theta$, the inequality (9) is satisfied:

$$Bel(B) \leq \sum\nolimits_{\theta \in B} p_\theta \leq Pl(B) \tag{9}$$

This inequality can be proved according to the properties of the upper and lower probability.

Probability distributions $\langle p_\theta | \theta \in \Theta \rangle$ also must meet the usual requirements for probability distributions, i.e.

$$\begin{cases} 0 \leq p_\theta \leq 1, \forall \theta \in \Theta \\ \sum_{\theta \in \Theta} p_\theta = 1 \end{cases} \tag{10}$$

It can be taken for granted that there are several probability distributions $\{p_\theta | \theta \in \Theta\}$ consistent with the given belief function according to the relationships defined in (9) and (10). This is a multi-answer problem or one-to-many mapping relation. As referred above, the probability is used for decision, so the uncertainty seemingly should be as little as possible. We can select one probability distribution from all the consistent alternatives according to the uncertainty minimization criterion and use the corresponding probability distribution as the result of the probability transformation.

The Shannon entropy is used here to establish the objective function. The equations and inequalities in (9) and (10) are used to establish the constraints. The problem of probability transform of belief function here is converted to an optimization problem under constraints as follows:

$$\begin{aligned} &\underset{\{p_\theta | \theta \in \Theta\}}{\text{Min}} \left\{ -\sum_{\theta \in \Theta} p_\theta \log_2(p_\theta) \right\} \\ &s.t. \begin{cases} Bel(B) \leq \sum_{\theta \in B} p_\theta \leq Pl(B) \\ 0 \leq p_\theta \leq 1, \forall \theta \in \Theta \\ \sum_{\theta \in \Theta} p_\theta = 1 \end{cases} \end{aligned} \tag{11}$$

Given belief function (or the BPA, the plausibility), by solving (11), a probability distribution can be derived, which has least uncertainty measured by Shannon entropy and thus is seemingly more proper to be used in decision procedure.

It is clear that the problem of finding a minimum entropy probability distribution does not admit a unique solution in general. The optimization algorithm used is the Quasi-Newton followed by a global optimization algorithm [16] to alleviate the effect of the local extremum problem. Other intelligent optimization algorithms [17, 18] can also be used,such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), etc.

4 Analysis based on examples

At first, two numerical examples are first provided to illustrate some probability transformation approaches. To make the different approaches reviewed and proposed in this paper more comparable, the examples in [6, 12] are directly used here. The PIC is used to evaluate the probability transformation.

4.1 Example 1

For FOD $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, the corresponding BPA is as follows:

$m(\{\theta_1\}) = 0.16,\ m(\{\theta_2\}) = 0.14,\ m(\{\theta_3\}) = 0.01,$
$m(\{\theta_4\}) = 0.02,$
$m(\{\theta_1, \theta_2\}) = 0.20,\ m(\{\theta_1, \theta_3\}) = 0.09,$
$m(\{\theta_1, \theta_4\}) = 0.04,\ m(\{\theta_2, \theta_3\}) = 0.04,$
$m(\{\theta_2, \theta_4\}) = 0.02,\ m(\{\theta_3, \theta_4\}) = 0.01,$
$m(\{\theta_1, \theta_2, \theta_3\}) = 0.10,\ m(\{\theta_1, \theta_2, \theta_4\}) = 0.03,$
$m(\{\theta_1, \theta_3, \theta_4\}) = 0.03,\ m(\{\theta_2, \theta_3, \theta_4\}) = 0.03,$
$m(\Theta) = 0.08.$

The corresponding belief functions are calculated and listed as follows:

$Bel(\{\theta_1\}) = 0.16,\ Bel(\{\theta_2\}) = 0.14,$
$Bel(\{\theta_3\}) = 0.01,\ Bel(\{\theta_4\}) = 0.02,$
$Bel(\{\theta_1, \theta_2\}) = 0.50,\ Bel(\{\theta_1, \theta_3\}) = 0.26,$
$Bel(\{\theta_1, \theta_4\}) = 0.22,\ Bel(\{\theta_2, \theta_3\}) = 0.19,$
$Bel(\{\theta_2, \theta_4\}) = 0.18,\ Bel(\{\theta_3, \theta_4\}) = 0.04,$
$Bel(\{\theta_1, \theta_2, \theta_3\}) = 0.74,\ Bel(\{\theta_1, \theta_2, \theta_4\}) = 0.61,$
$Bel(\{\theta_1, \theta_3, \theta_4\}) = 0.36,\ Bel(\{\theta_2, \theta_3, \theta_4\}) = 0.27,$
$Bel(\Theta) = 1.00.$

The corresponding plausibility functions are calculated and listed as follows:

$Pl(\{\theta_1\}) = 0.73,\ Pl(\{\theta_2\}) = 0.64,\ Pl(\{\theta_3\}) = 0.39,$
$Pl(\{\theta_4\}) = 0.26,$
$Pl(\{\theta_1, \theta_2\}) = 0.96,\ Pl(\{\theta_1, \theta_3\}) = 0.82,$
$Pl(\{\theta_1, \theta_4\}) = 0.81,\ Pl(\{\theta_2, \theta_3\}) = 0.78,$
$Pl(\{\theta_2, \theta_4\}) = 0.74,\ Pl(\{\theta_3, \theta_4\}) = 0.50,$
$Pl(\{\theta_1, \theta_2, \theta_3\}) = 0.98,\ Pl(\{\theta_1, \theta_2, \theta_4\}) = 0.99,$
$Pl(\{\theta_1, \theta_3, \theta_4\}) = 0.86,\ Pl(\{\theta_2, \theta_3, \theta_4\}) = 0.84,$
$Pl(\Theta) = 1.00.$

Suppose the probability distribution as the unknown variables. Based on the plausibility functions and the belief functions, the constraints and the objective function can be established according to (11). The probability distribution can be derived based on the minimization. The results of some other probability transformation approaches are also calculated. All the results are listed in Table 1 (on the next page) to make the comparison between the approach proposed in this paper (denoted by Un_min) and other available approaches.

4.2 Example 2

For FOD $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, the corresponding BBA is as follows:

$m(\{\theta_1\}) = 0.05,\ m(\{\theta_2\}) = 0.00,\ m(\{\theta_3\}) = 0.00,$
$m(\{\theta_4\}) = 0.00,$

Table **1** Probability Transformation Results of Example 1 based on Different Approaches

	θ_1	θ_2	θ_3	θ_4	PIC
BetP [3]	0.3983	0.3433	0.1533	0.1050	0.0926
PraPl [8]	0.4021	0.3523	0.1394	0.1062	0.1007
PrPl [8]	0.4544	0.3609	0.1176	0.0671	0.1638
PrHyb [8]	0.4749	0.3749	0.0904	0.0598	0.2014
PrBel [8]	0.5176	0.4051	0.0303	0.0470	0.3100
FPT[11]	0.5176	0.4051	0.0303	0.0470	0.3100
DSmP_0[9]	0.5176	0.4051	0.0303	0.0470	0.3100
PrScP [10]	0.5403	0.3883	0.0316	0.0393	0.3247
PrBP1 [12]	0.5419	0.3998	0.0243	0.0340	0.3480
PrBP2 [12]	0.5578	0.3842	0.0226	0.0353	0.3529
PrBP3 [12]	0.0605	0.3391	0.0255	0.0309	0.3710
Un_min	0.7300	0.2300	0.0100	0.0300	0.4813

$m(\{\theta_1,\theta_2\}) = 0.39,\ m(\{\theta_1,\theta_3\}) = 0.19,$
$m(\{\theta_1,\theta_4\}) = 0.18, m(\{\theta_2,\theta_3\}) = 0.04,$
$m(\{\theta_2,\theta_4\}) = 0.02,\ m(\{\theta_3,\theta_4\}) = 0.01,$
$m(\{\theta_1,\theta_2,\theta_3\}) = 0.04,\ m(\{\theta_1,\theta_2,\theta_4\}) = 0.02,$
$m(\{\theta_1,\theta_3,\theta_4\}) = 0.03,\ m(\{\theta_2,\theta_3,\theta_4\}) = 0.03,$
$m(\Theta) = 0.00.$

The corresponding belief functions are calculated and listed as follows:

$Bel(\{\theta_1\}) = 0.05,\ Bel(\{\theta_2\}) = 0.00,$
$Bel(\{\theta_3\}) = 0.00,\ Bel(\{\theta_4\}) = 0.00,$
$Bel(\{\theta_1,\theta_2\}) = 0.44,\ Bel(\{\theta_1,\theta_3\}) = 0.24,$
$Bel(\{\theta_1,\theta_4\}) = 0.23,\ Bel(\{\theta_2,\theta_3\}) = 0.04,$
$Bel(\{\theta_2,\theta_4\}) = 0.02,\ Bel(\{\theta_3,\theta_4\}) = 0.01,$
$Bel(\{\theta_1,\theta_2,\theta_3\}) = 0.71,\ Bel(\{\theta_1,\theta_2,\theta_4\}) = 0.66,$
$Bel(\{\theta_1,\theta_3,\theta_4\}) = 0.46,\ Bel(\{\theta_2,\theta_3,\theta_4\}) = 0.10,$
$Bel(\Theta) = 1.00.$

The corresponding plausibility functions are calculated and listed as follows:

$Pl(\{\theta_1\}) = 0.90,\ Pl(\{\theta_2\}) = 0.54,\ Pl(\{\theta_3\}) = 0.34,$
$Pl(\{\theta_4\}) = 0.29,$
$Pl(\{\theta_1,\theta_2\}) = 0.99,\ Pl(\{\theta_1,\theta_3\}) = 0.98,$
$Pl(\{\theta_1,\theta_4\}) = 0.96,\ Pl(\{\theta_2,\theta_3\}) = 0.77,$
$Pl(\{\theta_2,\theta_4\}) = 0.76,\ Pl(\{\theta_3,\theta_4\}) = 0.56,$
$Pl(\{\theta_1,\theta_2,\theta_3\}) = 1.00,\ Pl(\{\theta_1,\theta_2,\theta_4\}) = 1.00,$
$Pl(\{\theta_1,\theta_3,\theta_4\}) = 1.00,\ Pl(\{\theta_2,\theta_3,\theta_4\}) = 0.95,$
$Pl(\Theta) = 1.00.$

Suppose the probability distribution as the unknown variables. Based on the plausibility functions and the belief functions, the constraints and the objective function can be established according to (11). The probability distribution can be derived based on the minimization. The results of some other probability transformation approaches are also calculated. All the results are listed in Table 2 to make the comparison between the approach proposed in this paper and other available approaches.

N/A in Table 2 means "Not available". DSmP_0 means the parameter ε in DSmP is 0.

Table **2** Probability Transformation Results of Example 2 based on Different Approaches

	θ_1	θ_2	θ_3	θ_4	PIC
PrBel [8]	N/A due to 0 value of singletons				
FPT[11]	N/A due to 0 value of singletons				
PrScP [10]	N/A due to 0 value of singletons				
PrBP1 [12]	N/A due to 0 value of singletons				
PraPl [8]	0.4630	0.2478	0.1561	0.1331	0.0907
BetP [3]	0.4600	0.2550	0.1533	0.1317	0.0910
PrPl [8]	0.6161	0.2160	0.0960	0.0719	0.2471
PrBP2 [12]	0.6255	0.2109	0.0936	0.0700	0.2572
PrHyb [8]	0.6368	0.2047	0.0909	0.0677	0.2698
DSmP_0[9]	0.5162	0.4043	0.0319	0.0477	0.3058
PrBP3 [12]	0.8823	0.0830	0.0233	0.0114	0.5449
Un_min	0.9000	0.0900	0.0000	0.0100	0.7420

Based on the experimental results listed in Table 1 and Table 2, it can be concluded that the probability derived based on the proposed approach (denoted by Un_min) has significantly lower uncertainty when compared with the other probability transformation approaches. The difference among all the propositions can be further enlarged, which is seemingly helpful for the more consolidated and reliable decision.

Important remark: In fact, there exist fatal deficiencies in the probability transformation based uncertainty minimization, which are illustrated in following examples.

4.3 Example 3

The FOD and BPA are as follows [4]:

$\Theta = \{\theta_1,\theta_2\},\ m(\{\theta_1\}) = 0.3,$
$m(\{\theta_2\}) = 0.1,\ , m(\{\theta_1,\theta_2\}) = 0.6$

Based on different approaches, the experimental results are derived as listed in Table 3

Table **3** Probability Transformation Results of Example 3 based on Different Approaches

	θ_1	θ_2	PIC
BetP	0.6000	0.4000	0.0291
PrPl	0.6375	0.3625	0.0553
PraPl	0.6375	0.3625	0.0553
PrHyb	0.6825	0.3175	0.0984
DSmP_0.001	0.7492	0.2508	0.1875
PrBel	0.7500	0.2500	0.1887
DSmP_0	0.7500	0.2500	0.1887
Un_min	0.9000	0.1000	0.5310

DSmP_0 means the parameter ε in DSmP is 0 and DSmP_0.001 means the parameter ε in DSmP is 0.001.

Is the probability transformation based on PIC maximization (i.e. entropy minimization) rational ?

It can be observed, in our very simple example 3, that all the mass of belief 0.6 committed $\{\theta_1,\theta_2\}$ is actually redistributed only to the singleton $\{\theta_1\}$

using the Un_min transformation in order to get the maximum of PIC.

A deeper analysis shows that with Un_min transformation, the mass of belief $m\{\theta_1, \theta_2\} > 0$ is always fully distributed back to $\{\theta_1\}$ as soon as $m(\{\theta_1\}) > m(\{\theta_2\})$ in order to obtain the maximum of PIC (i.e. the minimum of entropy). Even in very particular situations where the difference between masses of singletons is very small like in the following example:

$\Theta = \{\theta_1, \theta_2\},\ m(\{\theta_1\}) = 0.1000001,$
$m(\{\theta_2\}) = 0.1,\ m(\{\theta_1, \theta_2\}) = 0.7999999.$

This previous modified example shows that the probability obtained from the minimum entropy principle yields a counter-intuitive result, because $m(\{\theta_1\})$ is almost the same as $m(\{\theta_2\})$ and so there is no solid reason to obtain a very high probability for θ_1 and a small probability for θ_2. Therefore, the decision based on the result derived from Un_min transformation is too risky. Sometimes uncertainty can be useful, and sometimes it is better to not take a decision than to take the wrong decision. So the criterion of uncertainty minimization is not sufficient for evaluating the quality/efficiency of a probability transformation. There are also other problems in the probability transformation based on uncertainty minimization principle, which are illustrated in our next example.

4.4 Example 4

The FOD and BPA are as follows: $\Theta = \{\theta_1, \theta_2, \theta_3\}$, with ,

$$m(\{\theta_1, \theta_2\}) = m(\{\theta_2, \theta_3\}) = m(\{\theta_1, \theta_3\}) = 1/3.$$

Using the probability transformation based on uncertainty minimization, we can derive six different probability distributions yielding the same minimal entropy, which are listed as follows:

$$\begin{array}{lll}
P(\{\theta_1\}) = 1/3, & P(\{\theta_2\}) = 2/3, & P(\{\theta_3\}) = 0; \\
P(\{\theta_1\}) = 1/3, & P(\{\theta_2\}) = 0, & P(\{\theta_3\}) = 2/3; \\
P(\{\theta_1\}) = 0, & P(\{\theta_2\}) = 1/3, & P(\{\theta_3\}) = 2/3; \\
P(\{\theta_1\}) = 0, & P(\{\theta_2\}) = 2/3, & P(\{\theta_3\}) = 1/3; \\
P(\{\theta_1\}) = 2/3, & P(\{\theta_2\}) = 1/3, & P(\{\theta_3\}) = 0; \\
P(\{\theta_1\}) = 2/3, & P(\{\theta_2\}) = 0, & P(\{\theta_3\}) = 1/3.
\end{array}$$

It is clear that the problem of finding a probability distribution with minimal entropy does not admit a unique solution in general. So if we use the probability transformation based on uncertainty minimization, there might exist several probability distributions derived as illustrated in this Example 4. How to choose a unique one? In Example 4, depending on the choice of the admissible probability distribution, the decision results derived are totally different which is a serious problem for decision-making support.

From our analysis, it can be concluded that the maximization of PIC criteria (or equivalently the minimization of Shannon entropy) is not sufficient for evaluating the quality of a probability transformation and other criteria have to be found to give more acceptable probability distribution from belief functions. The search for new criteria for developing new transformations is a very open and challenging problem. Until finding new better probability transformation, we suggest to use DSmP as one of the most useful probability transformation. Based on the experimental results shown in Examples 1–3, we see that the DSmP can always be computed and generate a probability distribution with less uncertainty and it is also not too risky, i.e. DSmP can achieve a better tradeoff between a high PIC value (i.e. low uncertainty) and the risk in decision-making.

5 Conclusion

Probability transformation of belief function can be considered as a probabilistic approximation of belief assignment, which aims to gain more reliable decision results. In this paper, we focus on the evaluation criteria of the probability transformation function. Experimental results based on numerical examples show that the maximization of PIC criteria proposed by Sudano is insufficient for evaluating the quality of a probability transformation. More rational criteria have to be found and to better justify the use of a probability transformation with respect to another one.

All the current probability transformations developed so far redistribute the mass of partial ignorances to the belief of singletons included in it. The redistribution is based either only on the cardinality of partial ignorances, or eventually also on a proportionalization using the masses of singletons involved in partial ignorances. However when the mass of a singleton involved in a partial ignorance is zero, some probability transformations, like Cuzzolin's transformation by example, do not work at all and that's why the ε parameter has been introduced in DSmP transformation to make it working in all cases. In future, we plan to develop a more comprehensive and rational criterion, which can take both the risk and the uncertainty degree into consideration, to evaluate the quality of a probability transformation and to find an optimal probability distribution from any basic belief assignment.

Acknowledgment

This work is supported by Grant for State Key Program for Basic Research of China (973) No.2007CB311006 and the Research Fund of Shaanxi Key Laboratory of Electronic Information System Integration (No.200910A).

References

[1] A. P. Dempster, "Upper and lower probabilities induced by a multi-valued mapping", *J. Annual Math Statist*, Vol 38, No. 4, pp. 325–339, 1967.

[2] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, 1976.

[3] P. Smets, R. Kennes, "The transferable belief model", *Artificial Intelligence*, Vol 66, No. 2, pp. 191-234, 1994.

[4] F. Smarandache, J. Dezert (Editors), *Applications and Advances of DSmT for Information Fusion*, Vol 3, American Research Press, 2009. http://www.gallup.unm.edu/˜smarandache/DSmT-book3.pdf.

[5] J. Sudano, "Pignistic probability transforms for mixes of low- and high-probability events", *Proc. of International Conference on Information Fusion 2001*, Montreal, Canada, August 2001, TUB3, pp. 23-27.

[6] J. Sudano,"The system probability information content (PIC) relationship to contributing components, combining independent multisource beliefs, hybrid and pedigree pignistic probabilities". *Proc. of International Conference on Information Fusion 2002*, Annapolis, July 2002. Vol 2, pp. 1277-1283.

[7] J. Sudano, "Equivalence between belief theories and naive Bayesian fusion for systems with independent evidential data - Part I, The theory", *Proc. of International Conference on Information Fusion 2003*, Cairns, Australia, July 2003, Vol 2, pp. 1239-1243.

[8] J. Sudano, "Yet another paradigm illustrating evidence fusion (YAPIEF)", *Proc. of International Conference on Information Fusion 2006*, Florence, July 2006, pp. 1-7.

[9] J. Dezert, F. Smarandache, "A new probabilistic transformation of belief mass assignment", *International Conference on Information Fusion 2008*, Germany, Cologne, June 30^{th} – July 3rd 2008, pp. 1-8.

[10] J. Sudano, "Belief fusion, pignistic probabilities, and information content in fusing tracking attributes", *Radar Conference 2004*, April 26th-29th 2004, pp. 218 – 224.

[11] Y. Deng, W. Jiang, Q. Li, "Probability transformation in transferable belief model based on fractal theory (in Chinese)", *Proc. of Chinese Conference on Information Fusion 2009*, Yantai, China, Nov. 2009, pp. 10-13, 79.

[12] W. Pan, H. J. Yang, "New methods of transforming belief functions to pignistic probability functions in evidence theory", *Proc. of International Workshop on Intelligent Systems and Applications*, Wuhan, China, May 2009, pp. 1-5.

[13] F. Cuzzolin, "On the properties of the Intersection probability", submitted to the *Annals of Mathematics and AI*, Feb. 2007.

[14] W. H. Lv, "Decision-making rules based on belief interval with D-S evidence theory, (book section)", *Fuzzy Information and Engineering*, pp. 619-627, 2007.

[15] P. Smets, "Decision making in the TBM: the necessity of the pignistic transformation", *International Journal of Approximate Reasoning*, Vol 38, pp. 133–147, 2005.

[16] R. Horst, P. M. Pardalos, N. V. Thoai, *Introduction to Global Optimization (2nd Edition)*, Kluwer Academic Publishers, New York, 2000.

[17] D. T. Pham, *Intelligent Optimisation Techniques (1st Edition)*, Springer-Verlag, London, 2000.

[18] J. Kennedy, R. C. Eberhart, Y. H. Shi, *Swarm Intelligence*, Morgan Kaufmann Publisher, Academic Press, USA, 2001

Algebraic Generalization of Venn Diagram

Florentin Smarandache

Abstract.

It is easy to deal with a Venn Diagram for $1 \leq n \leq 3$ sets. When n gets larger, the picture becomes more complicated, that's why we thought at the following codification. That's why we propose an easy and systematic algebraic way of dealing with the representation of intersections and unions of many sets.

Introduction.

Let's first consider $1 \leq n \leq 9$, and the sets $S_1, S_2, \ldots, S_n$.
Then one gets 2^n-1 disjoint parts resulted from the intersections of these n sets. Each part is encoded with decimal positive integers specifying only the sets it belongs to. Thus:
part 1 means the part that belongs to S_1 (set 1) only, part 2 means the part that belongs to S_2 only, ..., part n means the part that belongs to set S_n only.
Similarly, part 12 means that part which belongs to S_1 and S_2 only, i.e. to $S_1 \cap S_2$ only.
Also, for example part 1237 means the part that belongs to the sets S_1, S_2, S_3, and S_7 only, i.e. to the intersection $S_1 \cap S_2 \cap S_3 \cap S_7$ only. And so on. This will help to the construction of a base formed by all these disjoint parts, and implementation in a computer program of each set from the power set $\mathcal{P}(S_1 \quad S_2 \quad \ldots \quad S_n)$ using a binary number.
The sets $S_1, S_2, \ldots, S_n$, are intersected in all possible ways in a Venn diagram. Let $1 \leq k \leq n$ be an integer. Let's denote by: $i_1 i_2 \ldots i_k$ the Venn diagram region/part that belongs to the sets S_{i1} and S_{i2} and … and S_{ik} only, for all k and all n. The part which is outside of all sets (i.e. the complement of the union of all sets) is noted by 0 (zero). Each Venn diagram will have 2^n disjoint parts, and each such disjoint part (except the above part 0) will be formed by combinations of k numbers from the numbers: 1, 2, 3, …, n.

Example.

Let see an example for n = 3, and the sets S_1, S_2, and S_3.

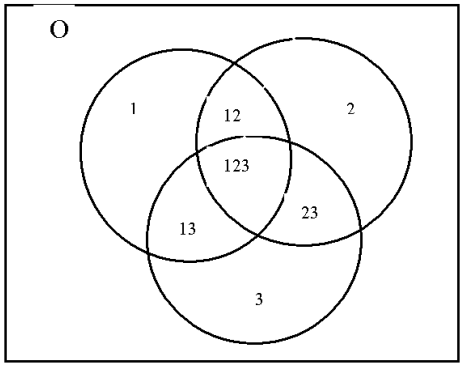


Fig. 1.

Unions and Intersections of Sets.

This codification is user friendly in algebraically doing unions and intersections in a simple way.
Union of sets S_a, S_b, …, S_v is formed by all disjoint parts that have in their index either the number a, or the number b, …, or the number v.
While intersection of S_a, S_b, …, S_v is formed by all disjoint parts that have in their index all numbers a, b, …, v.
For n = 3 and the above diagram:
$S_1 \cup S_{23}$ = {1, 12, 13, 23, 123}, i.e. all disjoint parts that include in their indexes either the digit 1, or the digits 23;
and $S_1 \cap S_2$ = {12, 123}, i.e. all disjoint parts that have in their index the digits 12.

Remarks.

When $n \geq 10$, one uses one space in between numbers: for example, if we want to represent the disjoint part which is the intersection of S_3, S_{10}, and S_{27} only, we use the notation [3 10 27], with blanks in between the set indexes.
Depending on preferences, one can use other character different from the blank in between numbers, or one can use the numeration system in base n+1, so each number/index will be represented by a unique character.

References:

[1] J. Dezert, F. Smarandache, *An introduction to DSmT*, in <Advances and Applications of DSmT for Information Fusion>, ARPress, Vol. 3, pp. 3-73, 2009.
[2] J. Dezert & F. Smarandache *On the generation of hyper-powersets for the DSmT*, Proc. Fusion 2003 Conf., Cairns, Australia.
[3] A. Martin, *Implementing general belief function framework with a practical codification for low complexity*, in <Advances and Applications of DSmT for Information Fusion>, Vol. 3, pp. 217-273, 2009;
and in http://arxiv.org/PS_cache/arxiv/pdf/0807/0807.3483v1.pdf .
[4] N. J. A. Sloane, <Encyclopedia of Integer Sequences>, *Sequence A082185. Smarandache's Codification used in computer programming*, http://www.research.att.com/~njas/sequences/A082185 .
[5] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion*, Vol. 1, Am. Res. Press, 2004, pp. 42-46.

Importance of Sources using the Repeated Fusion Method and the Proportional Conflict Redistribution Rules #5 and #6

Florentin Smarandache
Jean Dezert

Originally published as a scientific note 2010 in HAL archive. https://hal.archives-ouvertes.fr/hal-00471839. Printed with permission.

Abstract.
We present in this paper some examples of how to compute by hand the PCR5 fusion rule for three sources, so the reader will better understand its mechanism.
We also take into consideration the importance of sources, which is different from the classical discounting of sources.

1. Introduction.

Discounting of Sources.
Discounting a source $m_1(.)$ with the coefficient $0 \leq \alpha \leq 1$ and a source $m_2(.)$ with a coefficient $0 \leq \beta \leq 1$ (because we are not very confident in them), means to adjust them to $m_1'(.)$ and $m_2'(.)$ such that:
$m_1'(A) = \alpha \cdot m_1(A)$ for $A \neq \Theta$ (total ignorance), and $m_1'(\Theta) = \alpha \cdot m_1(\Theta) + 1 - \alpha$,
and $m_2'(A) = \beta \cdot m_2(A)$ for $A \neq \Theta$ (total ignorance), and $m_2'(\Theta) = \beta \cdot m_2(\Theta) + 1 - \beta$.

Importance of Sources using Repeated Fusion.
But if a source is more important than another one (since a such source comes from a more important person with a decision power, let's say an executive director), for example if source $m2(.)$ is twice more important than source $m_1(.)$, then we can combine $m_1(.)$ with $m_2(.)$ and with $m_2(.)$, so we repeated $m_2(.)$ twice. Doing this procedure, the source which is repeated (combined) more times than another source attracts the result towards its masses – see an example below.
Jean Dezert has criticized this method since if a source is repeated say 4 times and other source is repeated *6* times, then combining *4* times $m_1(.)$ with 6 times $m_2(.)$ will give a result different from combining *2* times $m_1(.)$ with *3* times $m_2(.)$, although $4/6 = 2/3$. In order to avoid this, we take the simplified fraction n/p, where $gcd(n, p) = 1$, where *gcd* is the greatest common divisor of the natural numbers n and p.
This method is still controversial since after a large number of combining n times $m_1(.)$ with p times $m_2(.)$ for $n+p$ sufficiently large, the result is not much different from a previous one which combines n_1 times $m_1(.)$ with p_1 times $m_2(.)$ for n_1+p_1 sufficiently large but a little less than $n+p$, so the method is not well responding for large numbers.

A more efficacy method of importance of sources consists in taking into consideration the discounting on the empty set and then the normalization (see especially paper [1] and also[2]).

2. Using m_{PCR5} for 3 Sources.

Example calculated by hand for combining three sources using *PCR5* fusion rule.

Let's say that $m_2(.)$ is 2 times more important than $m_1(.)$; therefore we fusion $m_1(.)$, $m_2(.)$, $m_2(.)$.

	A	B	$A \cup B$	$A \cap B = \Phi$
m_1	0.1	0.7	0.2	
m_2	0.4	0.1	0.5	
m_2	0.4	0.1	0.5	
m_{122}	0.193	0.274	0.050	0.483

$$\begin{cases} \dfrac{x_{1A}}{0.1} = \dfrac{y_{1B}}{0.7} = \dfrac{z_{2A\cup B}}{0.5} = \dfrac{0.005}{0.7} = \dfrac{0.05}{7} \\ x_{1A} = 0.000714 \\ y_{1B} = 0.000714 \\ z_{2A\cup B} = 0.003572 \end{cases}$$

$$\begin{cases} \dfrac{x_{2A}}{0.4} = \dfrac{y_{2B}}{0.7} = \dfrac{z_{2A\cup B}}{0.5} = \dfrac{0.14}{1.6} = \dfrac{0.07}{0.8} = \dfrac{0.7}{8} \\ x_{2A} = 0.035000 \\ y_{2B} = 0.061250 \\ z_{2A\cup B} = 0.043750 \end{cases}$$

$$\begin{cases} \dfrac{x_{3A}}{0.4} = \dfrac{y_{3B}}{0.1} = \dfrac{z_{3A\cup B}}{0.2} = \dfrac{0.008}{0.7} = \dfrac{0.08}{7} \\ x_{3A} \cong 0.004571 \\ y_{3B} \cong 0.001143 \\ z_{3A\cup B} \cong 0.002286 \end{cases}$$

$$\begin{cases} \dfrac{x_{4A}}{0.4} = \dfrac{y_{4B}}{0.1} = \dfrac{z_{4A\cup B}}{0.2} = \dfrac{(0.4)(0.1)(0.2)}{0.7} = \dfrac{0.008}{0.7} = \dfrac{0.08}{7} \\ x_{4A} \cong 0.004571 \\ y_{4B} \cong 0.001143 \\ z_{4A\cup B} \cong 0.002286 \end{cases}$$

$$\begin{cases} \dfrac{x_{5A}}{0.4} = \dfrac{y_{5B}}{0.7} = \dfrac{z_{5A\cup B}}{0.5} = \dfrac{0.14}{1.6} = \dfrac{1.4}{16} \\ x_{5A} \cong 0.035000 \\ y_{5B} \cong 0.061250 \\ z_{5A\cup B} \cong 0.043750 \end{cases}$$

$$\begin{cases} \dfrac{x_{6A}}{0.1} = \dfrac{y_{6B}}{0.1} = \dfrac{z_{6A\cup B}}{0.5} = \dfrac{0.005}{0.7} = \dfrac{0.05}{7} \\ x_{6A} \cong 0.000714 \\ y_{6B} \cong 0.000714 \\ z_{6A\cup B} \cong 0.003572 \end{cases}$$

$$\begin{cases} \dfrac{x_{7A}}{0.1} = \dfrac{y_{7B}}{(0.1)(0.1)} = \dfrac{(0.1)(0.1)(0.1)}{0.1+0.01} = \dfrac{0.001}{0.11} \\ x_{6A} \cong 0.000909 \\ y_{6B} \cong 0.000091 \end{cases}$$

$$\begin{cases} \dfrac{x_{8A}}{0.4} = \dfrac{y_{8B}}{(0.7)(0.1)} = \dfrac{(0.4)(0.7)(0.1)}{0.1+0.01} = \dfrac{0.028}{0.47} = \dfrac{2.8}{47} \\ x_{8A} \cong 0.023830 \\ y_{8B} \cong 0.004170 \end{cases}$$

$$x_{9A} = x_{8A} \cong 0.023830$$
$$y_{9B} = y_{8B} \cong 0.004170$$

$$\begin{cases} \dfrac{x_{10A}}{(0.1)(0.4)} = \dfrac{y_{10B}}{0.1} = \dfrac{(0.1)(0.4)(0.1)}{0.04+0.1} = \dfrac{0.004}{0.14} = \dfrac{0.4}{14} = \dfrac{0.2}{7} \\ x_{10A} \cong 0.001143 \\ y_{8B} \cong 0.002857 \end{cases}$$

$$x_{11A} = x_{10A} \cong 0.001143$$
$$y_{11B} = y_{10B} \cong 0.002857$$

$$\begin{cases} \dfrac{x_{12A}}{(0.1)(0.4)} = \dfrac{y_{12B}}{0.1} = \dfrac{(0.4)(0.4)(0.7)}{0.16+0.7} = \dfrac{0.112}{0.86} = \dfrac{11.2}{86} \\ x_{12A} \cong 0.020837 \\ y_{12B} \cong 0.091163 \end{cases}$$

	A	B	A∪B
m_{122}^{PCR5}	0.345262	0.505522	0.149216

If we didn't double $m_2(.)$ in the fusion rule, we'd get a different result.
Let's suppose we only fusion $m_1(.)$ with $m_2(.)$:

	A	B	A∪B	A∩B=Φ
m_1	0.1	0.7	0.2	
m_2	0.4	0.1	0.5	
m_{12}	0.17	0.44	0.10	0.29
m_{12}^{PCR5}	0.322	0.668	0.100	0

And now we compare the fusion results:

	A	B	A∪B	
m_{122}^{PCR5}	0.345	0.506	0.149	- *three sources*(second − source − doubled)*; importance of sources considered;*
m_{12}^{PCR5}	0.322	0.668	0.100	- *two sources; importance of sources not considered*.

The more times we repeat $m_2(.)$ the closer $m_{12...2}^{PCR5}(A) \rightarrow m_2(A)=0.4$, $m_{12...2}^{PCR5}(B) \rightarrow m_2(B)=0.1$, and $m_{12...2}^{PCR5}(A \cup B) \rightarrow m_2(A \cup B)=0.5$. Therefore, doubling, tripling, etc. a source, the mass of each element in the frame of discernment tends towards the mass value of that element in the repeated source (since that source is considered to have more importance than the others).

For the readers who want to do the previous calculation with a computer, here it is the m_{PCR5}
Formula for 3 Sources:

$$m_{PCR5}(A) = m_{123} + \sum_{\substack{X,Y \in G^{\Theta} \\ A \neq X \neq Y \neq A \\ A \cap X \cap Y = \Phi}} \left(\frac{m_1(A)^2 m_2(X) m_3(Y)}{m_1(A)+m_2(X)+m_3(Y)} + \right.$$

$$\left. + \frac{m_1(Y) m_2(A)^2 m_3(X)}{m_1(Y)+m_2(A)+m_3(X)} + \frac{m_1(X) m_2(Y) m_3(A)^2}{m_1(X)+m_2(Y)+m_3(A)} \right) +$$

$$+\sum_{\substack{X\in G^{\Theta}\\ A\cap X=\Phi}}\left(\frac{m_1(A)^2 m_2(X)m_3(X)}{m_1(A)+m_2(X)+m_3(X)}+\frac{m_1(X)m_2(A)^2 m_3(X)}{m_1(X)+m_2(A)+m_3(X)}+\frac{m_1(X)m_2(X)m_3(A)^2}{m_1(X)+m_2(X)+m_3(A)}\right)+$$

$$+\sum_{\substack{X\in G^{\Theta}\\ A\cap X=\Phi}}\left(\frac{m_1(A)^2 m_2(A)^2 m_3(X)}{m_1(A)+m_2(A)+m_3(X)}+\frac{m_1(X)m_2(A)^2 m_3(A)^2}{m_1(X)+m_2(A)+m_3(A)}+\frac{m_1(A)^2 m_2(X)m_3(A)^2}{m_1(A)+m_2(X)+m_3(A)}\right)$$

3. Similarly, let's see the m_{PCR6} Formula for 3 Sources:

$$m_{PCR6}(A)=m_{123}+\sum_{\substack{X,Y\in G^{\Theta}\\ A\neq X\neq Y\neq A\\ A\cap X\cap Y=\Phi}}\left(\frac{m_1(A)^2 m_2(X)m_3(Y)}{m_1(A)+m_2(X)+m_3(Y)}+\right.$$

$$\left.+\frac{m_1(Y)m_2(A)^2 m_3(X)}{m_1(Y)+m_2(A)+m_3(X)}+\frac{m_1(X)m_2(Y)m_3(A)^2}{m_1(X)+m_2(Y)+m_3(A)}\right)+$$

$$+\sum_{\substack{X\in G^{\Theta}\\ A\cap X=\Phi}}\left(\frac{m_1(A)^2 m_2(X)m_3(X)}{m_1(A)+m_2(X)+m_3(X)}+\frac{m_1(X)m_2(A)^2 m_3(X)}{m_1(X)+m_2(A)+m_3(X)}+\frac{m_1(X)m_2(X)m_3(A)^2}{m_1(X)+m_2(X)+m_3(A)}\right)+$$

$$+\sum_{\substack{X\in G^{\Theta}\\ A\cap X=\Phi}}\left(\frac{m_1(A)^2 m_2(A)m_3(X)+m_1(A)m_2(A)^2 m_3(X)}{m_1(A)+m_2(A)+m_3(X)}+\right.$$

$$+\frac{m_1(X)m_2(A)^2 m_3(A)+m_1(X)m_2(A)m_3(A)^2}{m_1(X)+m_2(A)+m_3(A)}+$$

$$\left.+\frac{m_1(A)^2 m_2(X)m_3(A)+m_1(A)m_2(X)m_3(A)^2}{m_1(A)+m_2(X)+m_3(A)}\right)$$

4. A General Formula for $PCR6$ for $s\geq 2$ Sources.

$$m_{PCR6}(A)=m_{12\ldots s}+\sum_{\substack{X_1,X_2,\ldots,X_{s-1}\in G^{\Theta}\\ X_i\neq A,i\in\{1,2,\ldots,s-1\}\\ \left(\bigcap_{i=1}^{s-1}X_i\right)\cap A=\Phi}}\sum_{k=1}^{s-1}\sum_{(i_1,i_2,\ldots,i_s)\in P(1,2,\ldots,s)}\left[m_{i_1}(A)+m_{i_2}(A)+\ldots+m_{i_k}(A)\right]\cdot$$

$$\cdot\frac{m_{i_1}(A)m_{i_2}(A)\ldots m_{i_k}(A)m_{i_{k+1}}(X_1)\ldots m_{i_s}(X_{s-k})}{m_{i_1}(A)+m_{i_2}(A)+\ldots+m_{i_k}(A)+m_{i_{k+1}}(X_1)+\ldots+m_{i_s}(X_{s-k})}$$

where *P(1, 2, …, s)* is the set of all permutations of the elements *{1, 2, …, s}*.

It should be observed that $X_1, X_2, \ldots, X_{s-1}$ may be different from each other, or some of them equal and others different, etc.

We wrote this *PCR6* general formula in the style of *PCR5*, different from Arnaud Martin & Christophe Oswald's notations, but actually doing the same thing. In order not to complicate the formula of *PCR6*, we did not use more summations or products after the third Sigma.

As a particular case:

$$m_{PCR6}(A) = m_{123} + \sum_{\substack{X_1,X_2\in G^\Theta \\ X_1\neq A, X_2\neq A \\ X_1\cap X_1\cap A=\Phi}} \sum_{k=1}^{2} \sum_{(i_1,i_2,i_3)\in P(1,2,3)} \frac{\left[m_{i_1}(A)+\ldots+m_{i_k}(A)\right] m_{i_1}(A)\ldots m_{i_k}(A) m_{i_{k+1}}(X_1)\ldots m_{i_3}(X_2)}{m_{i_1}(A)+\ldots+m_{i_k}(A)+m_{i_{k+1}}(X_1)+\ldots+m_{i_3}(X_2)}$$

where $P(1,2,3)$ is the set of permutations of the elements $\{1,2,3\}$.

It should also be observed that X_1 may be different from or equal to X_2.

Conclusion.

The aim of this paper was to show how to manually compute *PCR5* for *3* sources on some examples, thus better understanding its essence. And also how to take into consideration the *importance of sources* doing the Repeated Fusion Method. We did not present the Method of Discounting to the Empty Set in order to emphasize the importance of sources, which is better than the first one, since the second method was the main topic of paper [2].

We also presented the *PCR5* formula for *3* sources (a particular case when *n=3*), and the general formula for *PCR6* in a different way but yet equivalent to Martin-Oswald's *PCR6* formula.

References:

1. Dezert J., Tacnet J.-M., Batton-Hubert M., Smarandache F., *Multi-criteria Decision Making Based on DSmT-AHP*, in Proceedings of Workshop on the Theory of Belief Functions, April 1-2, 2010, Brest, France (available at http://www.ensieta.fr/belief2010/).
2. Smarandache F., Dezert J., Tacnet J.-M., *Fusion of Sources of Evidence with Different Importances and Reliabilities*, submitted to Fusion 2010, International Conference, Edinburgh, U.K., July 2010.
3. Smarandache Florentin, Dezert Jean, Li Xinde, *DSm Field and Linear Algebra of Refined Labels (FLARL)*, in the book "Advances and Applications of DSmT for Information Fusion", Am. Res. Press, Rehoboth, USA, Chapter 2 (pages 75-84), 2009; online at: http://fs.gallup.unm.edu//DSmT-book3.pdf .
4. Smarandache F., Dezert J., *Advances and Applications of DSmT for Information Fusion*, Vols. 1-3, Am. Res. Press, Rehoboth, 2004, 2006, 2009; http://fs.gallup.unm.edu//DSmT.htm .

Evidence Supporting Measure of Similarity for Reducing the Complexity in Information Fusion

Xinde Li
Jean Dezert
Florentin Smarandache
Xinhan Huang

Originally published as Xinde L, Dezert J., Smarandache F., Huang X., *Evidence Supporting Measure of Similarity for Reducing the Complexity in Information Fusion*, Information Sciences 181 (2011) 1818–1835, and reprinted with permission.

Abstract

This paper proposes a new solution for reducing the number of sources of evidence to be combined in order to diminish the complexity of the fusion process required in some applications where the real-time constraint and strong computing resource limitation are of prime importance. The basic idea consists in selecting, among the whole set of sources of evidence, only the biggest subset of sources which are not too contradicting based on a criterion of Evidence Supporting Measure of Similarity (ESMS) in order to process solely the coherent information received. The ESMS criterion serves actually as a generic tool for outlier source identification and rejection. Since the ESMS between several belief functions can be defined using several distance measures, we browse the most common ones in this paper and we describe in detail the principle of our Generalized Fusion Machine (GFM). The last part of the paper shows the improvement of the performances of this new approach with respect to the classical one in a real-data based and real-time experiment for robot perception using sonar sensors.

Key words:
Information fusion; Belief function; Complexity reduction; Robot perception; DSmT; Measure of similarity; Distance; Lattice.

1. Introduction

Information fusion (IF) has gained more and more interest in the scientific community since the end of nineties because of the development of sophisticated multisensor and hybrid (involving human feedbacks in the loop) systems in many fields of applications (robotics, defense, security, medicine, etc.). IF appears through many scientific international conferences and workshops [11]. The main theories useful for information fusion are the Probability theory [16, 25] (and more recently the Imprecise Probability Theory [38]), the Possibility Theory [8] (based on Fuzzy Sets theory [42]), Neutrosophic Set Theory [15] and belief function theories, mainly Dempster-Shafer theory (DST) [29] and more recently Dezert-Smarandache theory (DSmT) [31, 32, 33].

In this work, we concentrate our attention on belief functions theories and specially on DSmT because of its ability to deal efficiently with uncertain, imprecise and conflicting quantitative and qualitative information. Basically, in DST, a basic belief assignment (bba) $m(.)$ is a mapping from the power set 2^Θ (see section 2.2 for details) of the frame of discernment Θ into $[0,1]$ such that

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{X \in 2^\Theta} m(X) = 1.$$

In DST, Θ represents the set of exclusive and exhaustive possibilities for the solution of the problem under consideration. In DSmT, Θ can be a set of possible non exclusive elements and the definition of bba is extended to the lattice structures of hyper-power set D^Θ, and to super-power set S^Θ in UFT (Unification of Fusion Theories) [30, 32], Chap. 8 - see also section 2.3 for a brief presentation and [7, 10] and [33] for definitions, details and examples. In general $m(.)$ is not a measure of probability, except in the case when its focal elements (i.e. the elements which have a strictly positive mass of belief) are singletons; in such case, $m(.)$ is called a Bayesian bba [29] which can be considered as a subjective probability measure. In belief function theories, the main information fusion problem consists in finding an efficient way for combining several sources of evidence s_1, s_2, ..., s_n characterized by their bba's $m_1(.)$, $m_2(.)$, ..., $m_n(.)$ assumed for simplicity here defined on the same fusion space, either 2^Θ, D^Θ, or S^Θ depending on the underlying model associated with the nature of the frame Θ. The difficulty in information fusion arises from the fact that the sources can be conflicting (i.e.

one source commits some belief in a proposition A whereas another source commits some belief in a proposition B but A and B are known to be truly exclusive ($A \cap B = \emptyset$)) and one needs a solution for dealing with conflicting information in the fusion process. In DST, Shafer proposes Dempster's rule of combination as the fusion operator for combining sources of evidence whereas in DSmT the recommended fusion operator is the PCR5 (Proportional Conflict redistribution rule # 5) rule of combination, see [29] and [33] for discussions and comparisons of these rules. PCR5 is more complex than Demspter's rule but it offers a better ability to deal with conflicting information.

Both rules however become intractable in some applications having only low computational capacities (as in some autonomous onboard systems by example) because their complexity increases drastically with the number n of sources to combine and/or with the size of the frame Θ, specially in the worst case (i.e. when a strict positive mass of belief is committed to all elements of the fusion space). To circumvent this problem, one has to play on both sides: 1) reducing the number of sources to combine and 2) reducing the size of the frame Θ. In this paper, we propose a solution only for reducing the number of sources to combine because we are not concerned in our application of robot perception by the second aspect since in this application our frame Θ has only two elements representing the emptiness or occupancy states of the grid cells of the perceived map of the environment. To expect good performances of such limited-resource fusion scheme, it seems natural to search and combine altogether only the sources which are coherent (which are not too conflicting) according to a given measure of similarity.

Such idea has been already investigated by several authors who have proposed some distance measures between two evidential sources in different fields of applications. For example, Tessem [35] in 1993 proposed the distance $d_{ij} = max_{\theta_l \in \Theta} |BetP_i(\theta_l) - BetP_j(\theta_l)|)$ according to the pignistic probability transform $BetP(.)$. In 1997, Bauer [1] introduced two other measures of error to take a decision based on pignistic probability distribution after approximation. In 1998, Zouhal and Denoeux [43] also introduced a distance based on mean square error between pignistic probability. In 1999, Petit-Renaud [26] has defined a measure directly on the power set of Θ and proposed an error criterion between two belief structures based on the generalized Hausdorff distance. In 2001, Jousselme et al. [14] proposed in DST framework

a new distance measure $d_{ij} = 1 - \frac{1}{\sqrt{2}}\sqrt{m_1^2 + m_2^2 - 2\langle m_1, m_2\rangle}$ between two basic belief assignments (bba's) for measuring their similarity (closeness). In 2006, Ristic and Smets [27, 28] have defined in the TBM (Transferable Belief Model) framework a TBM-distance between bba's to solve the association of uncertain combat ID declarations. These authors recall also the Bhattacharya distance $d_{ij} = \sqrt{1 - \sum_{A\in\mathcal{F}_\rangle} \sum_{B\in\mathcal{F}_|} \sqrt{m_i(A)m_j(B)}}$ between two bba's. In 2006 also, Diaz et al. [6] proposed a new measure of similarity between bba's based on Tversky's similarity measure [37]. Note that in belief function theories, the direct use of classical measures used in Probability theory (say like Kullback Leibler (KL) distance [3]) cannot be applied directly because bba's are not probability measures in general.

In this paper, we develop an Evidence Support Measure of Similarity (ESMS) in a generalized fusion space according to different lattices [7, 10] for reducing the number of sources of evidence to combine and thus reducing the complexity of the computational burden. As shown in the next sections, we propose several possible measures of distance for ESMS and we compare their performances in our specific application of mobile robot perception. The purpose of this paper is not to select, nor to justify, the best measure of distance for ESMS but only to show the practical advantage of using the ESMS criteria as a generic tool for reducing the complexity of the fusion with keeping good performances for our application.

This paper is organized as follows. In section 2, we briefly recall the main paradigms for dealing with uncertain information. In section 3, we give a general mathematical definition of ESMS between two basic belief assignments and we establish some basic properties of ESMS. In section 4, we extend and present different possible ESMS functions (distance measures) fitting with the different mathematical paradigms listed in section 2. A comparison of the performances of five possible distances is made through a simple example in section 5. The simulation presented in section 6 shows in details how ESMS filter is used within GFM scheme. An application of ESMS filter in GFM for mobile robot perception with real-data (sonar sensors measurements) and in real-time is presented in section 7 to show the advantages of the approach proposed here. The conclusion is given in section 8.

2. The main paradigms for dealing with uncertainties

2.1. Probability Theory and Bayes' rule.

The (axiomatic) Probability Theory [16] is the most achieved theory for dealing with randomness. We will not present this theory in details since there exist dozens of very good classical books devoted to it, see for example [25]. We just recall that a random experiment is an experiment (action) whose result is uncertain before it is performed and a trial is a single performance of the random experiment. An outcome is the result of a trial and the sample space Θ is the set of all possible outcome of the random experiment. An event is the subset of the sample space Θ to which a probability measure can be assigned. Two events A_i and A_j are said exclusive (disjoint) if $A_i \cap A_j = \emptyset$, $\forall i \neq j$, where the empty set $\emptyset$ represents the impossible event. The sure event is the sample space Θ. The probability theory is based on Set Theory and the measure theory on sets. The following axioms have been identified as necessary and sufficient for probability $P(.)$ as a measure: Axiom 1) (nonnegativity) $0 \leq P(A) \leq 1$, Axiom 2) (unity) $P(\Theta) = 1$, and Axiom 3) (finite additivity[1]), if A_1, A_2, ..., A_n are disjoint events, then $P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} P(A_i)$. Events which are subsets of the sample space are put in one-to-one correspondence with propositions in belief fuction theory [29], pages 35-37 and that's why we use indifferently the terminology set, event or proposition in this paper. The probabilistic inference is (usually) carried out by Bayes' rule according to:

$$\forall B, P(B) > 0, \quad P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)} \qquad (1)$$

where the sample space Θ has been partitioned into exhaustive and exclusive events $A_1, A_2, \ldots, A_n$, i.e. such that $A_i \cap A_j = \emptyset, (i \neq j)$ and $A_1 \cup A_2 \cup \ldots \cup A_n = \Theta$; $P(.)$ is an a priori probability measure defined on Θ satisfying Kolmogorov's axioms. In Bayes formula, it is assumed that the denominator is strictly positive. A generalization of this rule has been proposed by Jeffrey [12, 13] for working in circumstances where the *parochialist* assumption is not a reasonable assumption, i.e. when $P(B|B) = 1$ is a fallacy, see [13, 22] for details and examples.

[1] Another axiom related to the countable additivity can be also considered as the fourth axiom of the probability theory.

Using the classical terminology adopted in belief function theories (DST and/or DSmT) and considering for example $\Theta = \{A, B\}$, a discrete probability measure $P(\cdot)$ can be interpreted as a specific Bayesian belief mass $m(.)$ such that

$$m(A) + m(B) = 1 \tag{2}$$

2.2. Dempster-Shafer Theory (DST)

In DST [29], the frame of discernment Θ of the fusion problem under consideration consists in a discrete finite set of n exhaustive and exclusive elementary hypotheses θ_i, i.e. $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$. This is called *Shafer's model* of the problem. Such model assumes that an ultimate refinement of the problem is possible, exists and is achievable, so that elements θ_i, $i = 1, 2, \ldots, n$ are well precisely defined and identified in such a way that we are sure that they are truly exclusive and exhaustive (closed-world assumption). The set of all subsets of Θ is called the power set of Θ and is denoted 2^Θ. Its cardinality is $2^{|\Theta|}$. Since 2^Θ is closed under $\cup$ and all θ_i, $i = 1, 2, \ldots, n$ are exclusive, it defines a Boolean algebra. All composite propositions built from elements of Θ with $\cup$ operator such that:

1) $\emptyset, \theta_1, \ldots, \theta_n \in 2^\Theta$;
2) If $A, B \in 2^\Theta$, then $A \cup B \in 2^\Theta$;
3) No other elements belong to 2^Θ, except those obtained by using rules 1) or 2).

Shafer defines a basic belief assignment (bba), also called mass function, as a mapping $m(.) : 2^\Theta \to [0, 1]$ satisfying $m(\emptyset) = 0$ and the normalization condition. Typically, when $\Theta = \{A, B\}$ and Shafer's model holds, in DST one works with $m(.)$ such that

$$m(A) + m(B) + m(A \cup B) = 1 \tag{3}$$

$m(A \cup B)$ allows us to commit some belief on the disjunction $A \cup B$ which represents the ignorance in choosing between A and B. From this very simple example, one sees clearly the ability of DST to offer a better modeling for a total ignorant/vacuous source of information by setting $m(A \cup B) = 1$, whereas in Probability Theory one would be forced to adopt the principle of insufficient reason (as known also as the principle of indifference) to justify taking $m(A) = m(B) = 1/2$ as default belief mass for representing a total ignorant body of evidence.

In DST framework, the combination of two belief assignments $m_1(.)$ and $m_2(.)$ is done using Dempster's rule of combination which can be seen as the normalized version of the conjunctive rule in order to remove the total conflicting mass and to get a proper normalized belief mass after the combination [29]. Dempster's rule is mathematically defined by $m(\emptyset) = 0$ and for $X \neq \emptyset$ by

$$m(X) = \frac{\sum\limits_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2)}{1 - \sum\limits_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2)} \tag{4}$$

Dempster's formula is defined if and only if the two sources of evidence are not fully conflicting; that is when $\sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2) \neq 1$.

2.3. Dezert-Smarandache Theory (DSmT)

In DSmT framework [31, 32, 33], the frame $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ is a finite set of n exhaustive elements which are not necessary exclusive. The principle of the third excluded middle and Shafer's model are refuted in DSmT (but can be introduced if needed depending on the model of the frame one wants to deal with), since for a wide class of fusion problems, the nature of hypotheses can be only vague and imprecise or crude approximation of the reality and none ultimate refinement is achievable. As a simple example, if we consider two suspects Peter (P) and Mary (M) in some criminal investigations, it may be possible that Peter has committed the crime alone, as well as Mary, or maybe Peter and Mary have committed the crime together. In that case, one has to consider the possibility for $P \cap M \neq \emptyset$ but there is no way to refine the original frame $\Theta = \{P, M\}$ into a finer one with exclusive finer elements say as $\Theta' = \{P \setminus (P \cap M), P \cap M, M \setminus (P \cap M)\}$ because there is no physical meaning and no possible occurrence of the atomic granules $P \setminus (P \cap M)$ and $M \setminus (P \cap M)$. In other words, the finer exclusive elements of the refined frame satisfying Shafer's model cannot always be well identified and precisely separated and they may have no sense at all. This is the main reason why DSmT allows as foundation the possibility to deal with non exclusive, partially overlapped or vague elements and refute Shafer's model and third excluded middle assumptions. DSmT proposes to work on a fusion

space defined by Dedekind's lattice also called *hyper-power set* D^Θ in DSmT.

The hyper-power set is defined as the set of all composite propositions built from elements of Θ with $\cap$ and $\cup$ operators such that [5]:

1) $\emptyset, \theta_1, \ldots, \theta_n \in D^\Theta$;
2) If $A, B \in D^\Theta$, then $A \cup B \in D^\Theta$ and $A \cap B \in D^\Theta$;
3) No other elements belong to D^Θ, except those obtained by using rules 1) or 2).

Following Shafer's idea, Dezert and Smarandache define a (generalized) basic belief assignment (or mass) as a mapping $m(.) : D^\Theta \rightarrow [0, 1]$ such that:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{X \in D^\Theta} m(X) = 1.$$

Typically, when $\Theta = \{A, B\}$ and Shafer's model doesn't hold, in DSmT one works with $m(.)$ such that

$$m(A) + m(B) + m(A \cup B) + m(A \cap B) = 1 \tag{5}$$

which appears actually as a direct and natural mathematical extension of (2) and (3).

Actually DSmT offers also the advantage to work with Shafer's model or with any hybrid model if some integrity constraints between elements of the frame are known to be true and must be taken into account in the fusion process. DSmT allows to solve static and/or dynamic[2] fusion problems in the same general mathematical framework. For notation convenience, one denotes by G^Θ the generalized fusion space or generalized power set including integrity constraints (i.e. exclusivity as well as possible non-existence restrictions between some elements of Θ), so that $G^\Theta = D^\Theta$ when no constraint enters in the model, or $G^\Theta = 2^\Theta$ when one wants to work with Shafer's model (see [31] for details and examples), or $G^\Theta = \Theta$ when working with probability model. If one wants to work with the space closed under union $\cup$, intersection $\cap$, and complementarity $\mathcal{C}$ operators, then $G^\Theta = S^\Theta$, i.e. the super-power set (see next section). A more general introduction of DSmT

[2] i.e. when the frame and/or its model change with time.

can be found in Chapter 1 of [33].

In DSmT, the fusion of two sources of evidences characterized by $m_1(.)$ and $m_2(.)$ is defined by $m_{PCR5}(\emptyset) = 0$ and $\forall X \in G^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{Y \in G^\Theta \\ X \cap Y = \emptyset}} [\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)}] \quad (6)$$

where all sets involved in formulas are in canonical form; $m_{12}(X) \equiv m_\cap(X) = \sum_{\substack{X_1, X_2 \in G^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2)$ corresponds to the conjunctive consensus on X between the $n = 2$ sources and where all denominators are different from zero. If a denominator is zero, that fraction is discarded. A general formula of PCR5 for the fusion of $n > 2$ sources has been proposed in [32].

2.4. Unification of Fusion Theory (UFT)

Recently Smarandache has proposed in [30, 32] an extension of DSmT by considering a super-power set S^Θ as the Boolean algebra on Θ, i.e. $S^\Theta = (\Theta, \cap, \cup, c(.))$. In other words, S^Θ is assumed to be closed under union $\cup$, intersection $\cap$, and complement $c(.)$ of sets respectively. With respect to the partial ordering relation, the inclusion $\subseteq$, the minimum element is the empty set $\emptyset$, and the maximal element is the total ignorance $I = \bigcup_{i=1}^n \theta_i$. Since it extends the power set space through the closed operation of $\cap$, $\cup$ and $c(.)$ operators, that is, UFT not only considers the non-exclusive situation among the elements, but also consider the exclusive, exhaustive, non-exhaustive situations, and even open and closed world. Typically, when $\Theta = \{A, B\}$, in UFT one works with $m(.)$ such that

$$m(A) + m(B) + m(A \cap B) + m(A \cup B) + m(c(A)) + m(c(B)) + m(c(A) \cup c(B)) = 1 \quad (7)$$

3. Evidence Support Measure of Similarity (ESMS)

Definition 3.1. *Let's consider a discrete and finite frame Θ and the fusion space G^Θ including integrity constraints of the model associated with Θ. The infinite set of basic belief assignments defined on G^Θ is denoted by $\mathfrak{m}_{G^\Theta}$. An Evidence Support Measure of Similarity (ESMS) of two (generalized) basic belief assignments $m_1(.)$ and $m_2(.)$ in $\mathfrak{m}_{G^\Theta}$ is the function$Sim(.,.) : \mathfrak{m}_{G^\Theta} \times \mathfrak{m}_{G^\Theta} \to [0, 1]$ satisfying the following conditions:*

1) Symmetry: $\forall m_1(.), m_2(.) \in \mathfrak{m}_{G^\Theta}, Sim(m_1, m_2) = Sim(m_2, m_1)$;
2) Consistency: $\forall m(.) \in \mathfrak{m}_{G^\Theta}, Sim(m, m) = 1$;
3) Non-negativity: $\forall m_1(.), m_2(.) \in \mathfrak{m}_{G^\Theta}, Sim(m_1, m_2) \geq 0$

We will say that $m_2(.)$ is more similar to $m_1(.)$ than $m_3(.)$ if and only if $Sim(m_1, m_2) \geq Sim(m_1, m_3)$. The maximum degree of similarity is naturally obtained when both bba's $m_1(.)$ and $m_2(.)$ coincide, which is expressed by consistency condition 2). The equality $Sim(m_1, m_2) = 0$ must be obtained when bba's have no focal elements in common, in particular whenever $m_1(.)$ is focused on $X \in G^\Theta$, which is denoted $m_1^X(.)$ and corresponds to $m_1(X) = 1$, and $m_2(.)$ is focused on $Y \in G^\Theta$, i.e. $m_2(.) = m_2^Y(.)$ such that $m_2(Y) = 1$, with $X \cap Y = \emptyset$.

Theorem 3.1. *For any bba* $m_1(.) \in \mathfrak{m}_{G^\Theta}$ *(which is a* $|G^\Theta|$*-dimensional vector) and any small positive real number* ϵ*, there exists at least one bba* $m_2(.) \in \mathfrak{m}_{G^\Theta}$ *for a given distance measure*[3] $d(.,.)$ *such that* $d(m_1, m_2) \leq \epsilon$.

Proof: Let's take $m_2(.) = m_1(.)$, then $d(m_1, m_2) = d(m_1, m_1) = d(m_2, m_2) = 0 < \epsilon$ which completes the proof.

Definition 3.2. *(Agreement of evidence) : If there exist two basic belief assignments* $m_1(.)$ *and* $m_2(.)$ *in* $\mathfrak{m}_{G^\Theta}$ *such that for some distance measure* $d(.,.)$*, one has* $d(m_1, m_2) \leq \epsilon$ *with* $\epsilon > 0$*, then* ϵ *is called the agreement of evidence supporting measure between* $m_1(.)$ *and* $m_2(.)$ *with respect to the chosen distance* $d(.,.)$*.* $m_1(.)$ *and* $m_2(.)$ *are said* ϵ*-consistent with respect to the distance* $d(.,.)$*.*

Theorem 3.2. *The smaller* $\epsilon > 0$ *is, the closer the distance* $d(m_1, m_2)$ *between* $m_1(.)$ *and* $m_2(.)$ *is, that is, the more similar or consistent* $m_1(.)$ *and* $m_2(.)$ *are.*

Proof: According to the Definition 3.2, if the evidence measure between $m_1(.)$ and $m_2(.)$ is ϵ-consistent, then $d(m_1, m_2) \leq \epsilon$. Let's take $\epsilon = 1 - Sim(m_1, m_2)$; when ϵ becomes smaller and smaller, $Sim(m_1, m_2)$ becomes greater and greater, according to the definition of ESMS and thus more similar or consistent $m_1(.)$ and $m_2(.)$ become. Finally, if $\epsilon = 1 - Sim(m_1, m_2) = 0$,

[3]Here we don't specify the distance measure and keep it only as a generic distance. Actually $d(.,.)$ can be any distance measure. In practice, the Euclidean distance is frequently used.

then $m_1(.)$ and $m_2(.)$ are totally consistent.

From the previous definitions and theorems, the ESMS appears as an interesting measure for evaluating the degree of similarity between two sources. We propose to use ESMS in a pre-processing/thresholding technique in order to reduce the complexity of the combination of sources of evidence by keeping in the fusion process only the sources which are ϵ-consistent. ϵ is actually a threshold parameter which has to be tuned by the system designer and which depends on the application and computational resources.

4. Several possible ESMS

In this section we propose several possibilities for choosing an ESMS function $Sim(.,.)$ satisfying theorem 3.1.

4.1. Euclidean ESMS function $Sim_E(m_1, m_2)$

Definition 4.1. *Let $\Theta = \{\theta_1, \ldots, \theta_n\}$ $(n > 1)$, $m_1(.)$ and $m_2(.)$ in $\mathfrak{m}_{G^\Theta}$, X_i the i-th (generic) element of G^Θ and $|G^\Theta|$ the cardinality of G^Θ. The following simple Euclidean ESMS function can be extended from [14]:*

$$Sim_E(m_1, m_2) = 1 - \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{|G^\Theta|} \left(m_1(X_i) - m_2(X_i)\right)^2} \tag{8}$$

The following theorem establishes that $Sim_E(m_1, m_2)$ is an ESMS function.

Theorem 4.1. *$Sim_E(m_1, m_2)$ defined in (8) is an ESMS function.*

Proof:

1) Let's prove that $Sim_E(m_1, m_2) \in [0, 1]$. If $Sim_E(m_1, m_2) > 1$, from (8) one would get $\frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^{|G^\Theta|} \left(m_1(X_i) - m_2(X_i)\right)^2} < 0$ which is impossible, so that $Sim_E(m_1, m_2) \leq 1$. Let's prove $Sim_E(m_1, m_2) \geq 0$ or equivalently from (8), $\sum_{i=1}^{|G^\Theta|} \left(m_1(X_i) - m_2(X_i)\right)^2 \leq 2$. This inequality is equivalent to $\sum_{i=1}^{|G^\Theta|} m_1(X_i)^2 + \sum_{i=1}^{|G^\Theta|} m_2(X_i)^2 \leq 2 + 2 \sum_{i=1}^{|G^\Theta|} m_1(X_i)$ $m_2(X_i)$. We denote it (i) for short. (i) always holds because one has $(\sum_{i=1}^{|G^\Theta|} m_1(X_i)^2 + \sum_{i=1}^{|G^\Theta|} m_2(X_i)^2) \leq ([\sum_{i=1}^{|G^\Theta|} m_1(X_i)]^2 + [\sum_{i=1}^{|G^\Theta|} m_2(X_i)]^2)$ and thus

$(\sum_{i=1}^{|G^\Theta|} m_1(X_i)^2 + \sum_{i=1}^{|G^\Theta|} m_2(X_i)^2) \leq 2$ because $[\sum_{i=1}^{|G^\Theta|} m_s(X_i)]^2 = 1$ for $s = 1, 2$ ($m_s(.)$ being normalized bba). Therefore inequality (i) holds and thus $Sim_E(m_1, m_2) \geq 0$.

2) It is easy to check that $Sim_E(m_1, m_2)$ satisfies the first condition of Definition 3.1.

3) If $m_1(.) = m_2(.)$, then $Sim_E(m_1, m_2) = 1$ because

$$\sum_{i=1}^{|G^\Theta|} (m_1(X_i) - m_2(X_i))^2 = 0.$$

Thus the second condition of Definition 3.1 is also satisfied.

4) Non-negativity has been proven above in the first part. Herein we use a particular case to show that Sim(m1, m2)=0, i.e. there exist m_1^X and m_2^Y for some $X, Y \in G^\Theta \setminus \{\emptyset\}$ such that $X \neq Y$, then according to (8), one gets $\sum_{i=1}^{|G^\Theta|} (m_1(X_i) - m_2(X_i))^2 = [m_1^X(X)]^2 + [m_2^Y(Y)]^2 = 2$ and thus one has $Sim_E(m_1^X, m_2^Y) = 1 - (\sqrt{2}/\sqrt{2}) = 0$, so that $Sim_E(.,.)$ verifies the third condition of Definition 3.1.

4.2. Jousselme ESMS function $Sim_J(m_1, m_2)$

Definition 4.2. *Let $m_1(.)$ and $m_2(.)$ be two basic belief assignments in $\mathfrak{m}_{G^\Theta}$ provided by the sources of evidence S_1 and S_2. Given a $|G^\Theta| \times |G^\Theta|$ assumed*[4] *positive definite matrix $\mathbf{D} = [D_{ij}]$, where $D_{ij} = |X_i \cap X_j|/|X_i \cup X_j|$, with $X_i, X_j \in G^\Theta$. Then, Jousselme ESMS function can be redefined from the Jousselme et al. measure [14]:*

$$Sim_J(m_1, m_2) = 1 - \frac{1}{\sqrt{2}} \sqrt{(m_1 - m_2)^T \mathbf{D} (m_1 - m_2)} \tag{9}$$

or equivalently

$$Sim_J(m_1, m_2) = 1 - \frac{1}{\sqrt{2}} \sqrt{m_1^2 + m_2^2 - 2\langle m_1, m_2 \rangle}$$

[4] Actually, Jousselme et al. in [14] did not prove that $\mathbf{D} = [D_{ij} = |X_i \cap X_j|/|X_i \cup X_j|]$ is truly a positive definite matrix. $\mathbf{D}$ is until now assumed to be positive definite. This is only a conjecture and proving it is not a trivial problem.

where $\langle m_1, m_2 \rangle$ is the scalar product defined as

$$\langle m_1, m_2 \rangle = \sum_{i=1}^{|G^\Theta|} \sum_{j=1}^{|G^\Theta|} D_{ij} m_1(X_i) m_2(X_j)$$

$X_i, X_j \in G^\Theta$, $i, j = 1, \ldots, s, |G^\Theta|$; $\|m\|^2$ represents the squared norm of the vector (bba) m, i.e. $\|m\|^2 = \langle m, m \rangle$.

Theorem 4.2. *$Sim_J(m_1, m_2)$ defined in formula (9) is an ESMS function.*

Proof:

1) Since the matrix $\mathbf{D}$ is conjectured to be a positive definite matrix, $Sim_J(m_1, m_2)$ satisfies the condition of symmetry.
2) If m_1 is equal to m_2, according to (9), one gets $Sim_J(m_1, m_2) = 1$. In other hand, if $Sim_J(m_1, m_2) = 1$, then the condition $m_1 = m_2$ holds. That is, the condition of consistency is satisfied.
3) According to (9), it can be drawn that $Sim_J(m_1, m_2) \leq Sim_E(m_1, m_2)$, and since the minimum value of $Sim_J(m_1, m_2)$ is zero, then $Sim_J(m_1, m_2)$ is non-negative.
4) According to the definition of $Sim_J(m_1, m_2)$, we can easily verify that $Sim_J(m_1, m_2)$ is a true distance measure between m_1 and m_2.

Actually $Sim_E(m_1, m_2)$ is nothing but a special case of $Sim_J(m_1, m_2)$ when taking $\mathbf{D}$ as the $|G^\Theta| \times |G^\Theta|$ identity matrix.

4.3. Ordered ESMS function $Sim_O(m_1, m_2)$

The definition of this (partial) ordered-based ESMS function is similar to $Sim_J(m_1, m_2)$ but instead of using Jousselme's matrix $\mathbf{D} = [D_{ij}]$, where $D_{ij} = |X_i \cap X_j|/|X_i \cup X_j|$, with $X_i, X_j \in G^\Theta$, we choose the DSm matrix $\mathbf{S} = [S_{ij}]$ where $S_{ij} = s(X_i \cap X_j)/s(X_i \cup X_j)$. Therefore, one has

$$Sim_O(m_1, m_2) = 1 - \frac{1}{\sqrt{2}} \sqrt{(m_1 - m_2)^T \mathbf{S} (m_1 - m_2)} \qquad (10)$$

The function $s(X)$ corresponds to the intrinsic informational content of the proposition X defined in details in [31] (Chap. 3) which is used for

partially ordering the elements of G^Θ. More precisely, $s(X)$ is the sum of the inverse of the length of the components of Smarandache's code[5] of X.

As a simple example, let's take $\Theta = \{\theta_1, \theta_2\}$ with free DSm model (i.e i.e. when all elements are non-exclusive two by two), then the partially[6] ordered hyper-power set G^Θ is given by $G^\Theta = \{\emptyset, \theta_1 \cap \theta_2, \theta_1, \theta_2, \theta_1 \cup \theta_2\}$ because $s(\emptyset) = 0$, $s(\theta_1 \cap \theta_2) = 1/2$, $s(\theta_1) = 1 + 1/2$, $s(\theta_2) = 1 + 1/2$ and $s(\theta_1 \cup \theta_2) = 1 + 1 + 1/2$ since Smarandache's codes of $\emptyset$, θ_1, θ_2, $\theta_1 \cap \theta_2$ and $\theta_1 \cup \theta_2$ are respectively given by $\{< . >\}$ (empty code), $\{< 1 >, < 12 >\}$, $\{< 2 >, < 12 >\}$, $\{< 12 >\}$ and $\{< 1 >, < 12 >, < 2 >\}$. The matrix $\mathbf{S}$ is defined by[7]

$$\mathbf{S} = \begin{bmatrix} \frac{s(\theta_1\cap\theta_2)}{s(\theta_1\cap\theta_2)} & \frac{s(\theta_1\cap\theta_2)}{s(\theta_1)} & \frac{s(\theta_1\cap\theta_2)}{s(\theta_2)} & \frac{s(\theta_1\cap\theta_2)}{s(\theta_1\cup\theta_2)} \\ \frac{s(\theta_1\cap\theta_2)}{s(\theta_1)} & \frac{s(\theta_1)}{s(\theta_1)} & \frac{s(\theta_1\cap\theta_2)}{s(\theta_1\cup\theta_2)} & \frac{s(\theta_1)}{s(\theta_1\cup\theta_2)} \\ \frac{s(\theta_1\cap\theta_2)}{s(\theta_2)} & \frac{s(\theta_1\cap\theta_2)}{s(\theta_1\cup\theta_2)} & \frac{s(\theta_2)}{s(\theta_2)} & \frac{s(\theta_2)}{s(\theta_1\cup\theta_2)} \\ \frac{s(\theta_1\cap\theta_2)}{s(\theta_1\cup\theta_2)} & \frac{s(\theta_1)}{s(\theta_1\cup\theta_2)} & \frac{s(\theta_2)}{s(\theta_1\cup\theta_2)} & \frac{s(\theta_1\cup X_2)}{s(\theta_1\cup\theta_2)} \end{bmatrix} = \begin{bmatrix} 1 & 1/3 & 1/3 & 1/5 \\ 1/3 & 1 & 1/5 & 3/5 \\ 1/3 & 1/5 & 1 & 3/5 \\ 1/5 & 3/5 & 3/5 & 1 \end{bmatrix}$$

It is easy to verify the positiveness of the matrix S by checking the positivity of all its eigenvalues which are $\lambda_1 = 0.800 > 0$, $\lambda_2 \approx 0.835 > 0$, $\lambda_3 \approx 0.205 > 0$ and $\lambda_4 \approx 2.160 > 0$. We have verified the positiveness of matrix $\mathbf{S}$ for $Card(\Theta) = n \leq 5$. Since a general proof of the positiveness of $\mathbf{D}$ and $\mathbf{S}$ seems difficult to obtain, we can only make a conjecture on the positiveness of $\mathbf{S}$ presently.

4.4. ESMS function $Sim_B(m_1, m_2)$

Another ESMS function based on Bhattacharya's distance is defined as follows:

Definition 4.3. *Let $m_1(.)$, $m_2(.)$ be two basic belief assignments in $\mathfrak{m}_{G^\Theta}$, the ESMS function $Sim_B(m_1, m_2)$ is defined by:*

[5]Smarandache code is a representation of disjoint parts of the Venn diagram of the frame Θ under consideration. This code depends of the model for Θ. For example, let's take $\Theta = \{\theta_1, \theta_2\}$. If $\theta_1 \cap \theta_2 = \emptyset$ (Shafer's model) is assumed, then the code of θ_1 is $< 1 >$, whereas if $\theta_1 \cap \theta_2 \neq \emptyset$ (free DSm model) is assumed, then the code of θ_1 will be $\{< 1 >, < 12 >\}$. The length of a component of a code is the number of characters between $<$ and $>$ in Smarandache's notation. For example, the length of component $< 12 >$ is 2. See [31], pp. 42–43 for details.

[6]This is a partial order since $s(\theta_1) = s(\theta_2)$.

[7]Actually, one works with $G^\Theta \setminus \{\emptyset\}$, and thus the column and row corresponding to the empty set do not enter in the definition of $\mathbf{S}$.

$$Sim_B(m_1, m_2) = 1 - \sqrt{1 - \sum_{X_i \in \mathcal{F}} \sqrt{m_1(X_i) m_2(X_i)}} \tag{11}$$

where $\mathcal{F}$ is the core of sources S_1 and S_2 , i.e. the set of elements of G^Θ having a positive belief mass: $\mathcal{F} = \{X \in G^\Theta | m_1(X) > 0$ or $m_2(X) > 0\}$.

Theorem 4.3. *$Sim_B(m_1, m_2)$ defined in formula (11) is an ESMS function.*

Proof:

1) Since $\sum_{X_i \in \mathcal{F}} \sqrt{m_1(X_i) m_2(X_i)} = \sum_{X_i \in \mathcal{F}} \sqrt{m_2(X_i) m_1(X_i)}$ then $Sim_B(m_1, m_2)$ satisfies the condition of symmetry.
2) if $m_1(.) = m_2(.)$, according to (11),

$$\sum_{X_i \in \mathcal{F}} \sqrt{m_1(X_i) m_2(X_i)} = \sum_{X_i \in \mathcal{F}} m_1(X_i) = 1$$

and therefore $Sim_B(m_1, m_1) = 1$. In other hand, if $Sim_B(m_1, m_2) = 1$, then the condition $m_1(.) = m_2(.)$ holds. That is, the condition of consistency is satisfied.
3) From the definition of bba, $\sum_{X_i \in \mathcal{F}} m_1(X_i) = 1$. Therefore,

$$\sum_{X_i \in \mathcal{F}} \sqrt{m_1(X_i) m_2(X_i)} \in [0, 1].$$

According to (11), it can be drawn that $Sim_B(m_1, m_2) \in [0, 1]$; that is, the minimum value of $Sim_B(m_1, m_2)$ is zero. Therefore, $Sim_B(m_1, m_2)$ is a non-negative measure.
4) According to the definition of $Sim_B(m_1, m_2)$, we can easily verify that $Sim_B(m_1, m_2)$ is a true distance measure between m_1 and m_2.

5. Comparison of ESMS functions

In this section we analyze the performances of the five ESMS functions aforementioned through a very simple example, where $\Theta = \{\theta_1, \theta_2, \theta_3\}$. We assume the free DSm model for Θ. In such case, $G^\Theta = D^\Theta$ has eighteen non-empty elements a_i, $i = 1, 2, \ldots, s, \ldots, 18$. G^Θ is closed under $\cap$ and $\cup$

operators according to Dedekind's lattice. Of course, we also may choose other theoretical frameworks in a similar way according to the appropriate model of G^Θ.

Let's assume that θ_2 is the true identity of the object under consideration. Its optimal belief assignment is denoted $m_2(.) \triangleq \{m_2(\theta_2) = 1, m_2(X) = 0 \text{ for } X \in G^\Theta \setminus \{\theta_2\}\}$. We perform a comparison of the four ESMS functions in order to show the evolution of the measure of similarity between $m_1(.)$ and $m_2(.)$ when $m_1(.)$ is varying from an uniform distributed bba to $m_2(.)$. More precisely, we start our simulation by choosing $m_1(.)$ with all elements in D^Θ uniformly distributed, i.e. $m_1(a_i) = 1/18$, for $i = 1, 2, \ldots, s, \ldots, 18$. Then, step by step we increase the mass of belief of θ_2 by a constant increment $\Delta = 0.01$ until reaching $m_1(\theta_2) = 1$. In the meantime the mass $m_1(X)$ of belief of all elements $X \neq \theta_2$ of G^Θ take value $[1 - m_1(\theta_2)]/17$ in order to work with a normalized bba $m_1()$. The basic belief mass committed to empty set is always zero, i.e. $m_1(\emptyset) = m_2(\emptyset) = 0$.

The degree of similarity of the four ESMS functions are plotted in Figure 1. The speed of convergence[8] of a similarity measure is characterized by the angle α of the slope of the curve at origin, or by its tangent. Based on this speed of convergence criterion, the analysis of the figure 1 yields the following remarks:

1) According to Figure 1., $\tan(\alpha_B) \approx 0.86$, $\tan(\alpha_E) \approx 0.68$, $\tan(\alpha_O) \approx 0.6$ and $\tan(\alpha_J) \approx 0.57$. $Sim_J(m_1, m_2)$ has the slowest convergence speed, then $Sim_O(m_1, m_2)$ takes second place.
2) $Sim_E(m_1, m_2)$ has a faster speed of convergence than $Sim_O(m_1, m_2)$ and $Sim_J(m_1, m_2)$ because it doesn't consider the intrinsic complexity of the elements in G^Θ.
3) The speed of convergence of $Sim_B(m_1, m_2)$ is the fastest. When $m_1(.)$ and $m_2(.)$ become very similar, $Sim_B(m_1, m_2)$ becomes very quickly close to 1. But if a small dissimilarity between $m_1()$ and $m_2()$ occurs, then $Sim_B(m_1, m_2)$ becomes very small which actually makes it very sensitive to small dissimilarity perturbations.

[8]Here the convergence speed refers to how much the global agreement degree (similarity) is between $m_1(.)$ and $m_2(.)$ with the continuous decrease of $m_1(\theta_2)$.

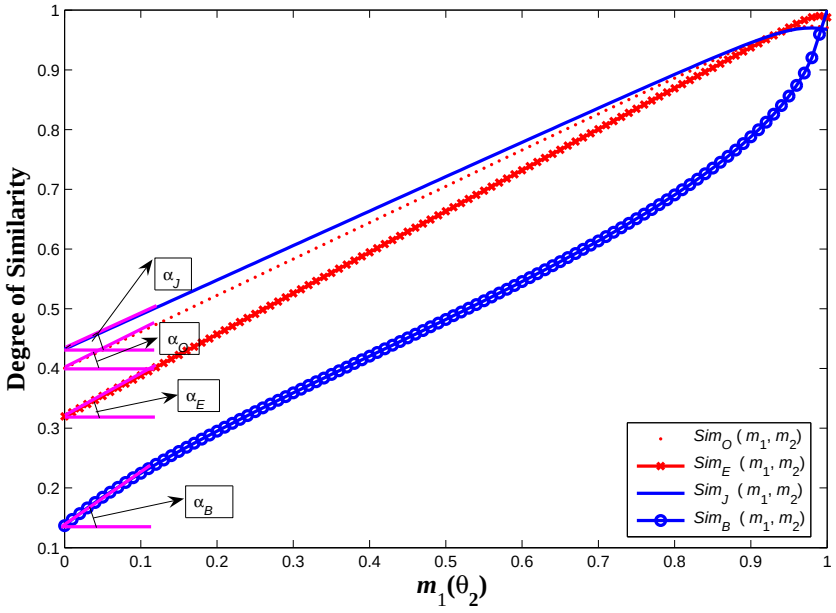


Figure 1: Comparison of performance among ESMS functions.

4) In summary, one sees that no definitive conclusion about the best choice among these four ESMS functions can be drawn in general, but if one considers as important the speed of convergence criterion of the dissimilarity/ difference between two evidential sources, $Sim_B(m_1, m_2)$ is the best choice, because it is very sensitive to such difference, whereas $Sim_J(m_1, m_2)$ is the worst choice with respect to such criterion.

6. Simulation results

We present a simulation result to show how the ESMS filter performs in generalized fusion machine (GFM), and its advantage. Let's take a 2D frame of discernment $\Theta = \{\theta_1, \theta_2\}$ and consider twenty equireliable sources of evidence according to the Table 1. We consider the free DSm model and the fusion space is the hyper-power set $D^\Theta = \{\theta_1, \theta_2, \theta_1 \cap \theta_2, \theta_1 \cup \theta_2\}$. $\mathbf{S}_{c1}$ denotes the barycentre of the front 10 belief masses, while $\mathbf{S}_{c2}$ denotes

the barycentre[9] of all belief masses . The measure of similarity based on Euclidean ESMS function defined in equation (8) has been used here, but any other measures of similarity could be used instead. In this example if we take 0.75 for the threshold value we see from the Table 1 and for the 10 front sources of evidences, that the measures of similarity of S_5 and S_{10} with respect to $\mathbf{S}_{c1}$ are lower than 0.75. Therefore, the sources S_5 and S_{10} will be discarded/filtered of the fusion process. If the threshold value is set to 0.8, then the sources S_5, S_{10}, S_4 and S_8 will be discarded. That is, the higher the given threshold is, the less the number of information sources through the filter is.

S	$m(\theta_1)$	$m(\theta_2)$	$m(\theta_1 \cap \theta_2)$	$m(\theta_1 \cup \theta_2)$	Sim_E
S_1	0.3	0.4	0.2	0.1	0.8735
S_2	0.3	0.2	0.4	0.1	0.800
S_3	0.4	0.1	0.2	0.3	0.8268
S_4	0.7	0.1	0.1	0.1	0.7592
S_5	0.1	0.8	0.1	0.0	0.5550
S_6	0.5	0.2	0.2	0.1	0.9106
S_7	0.4	0.3	0.1	0.2	0.9368
S_8	0.3	0.1	0.2	0.4	0.7592
S_9	0.4	0.5	0.1	0.0	0.8103
S_{10}	0.8	0.1	0.0	0.1	0.6806
$\mathbf{S}_{c1}$	0.42	0.28	0.16	0.14	1.0000
S_{11}	0.5	0.0	0.2	0.3	0.7569
S_{12}	0.2	0.6	0.1	0.1	0.7205
S_{13}	0.4	0.3	0.2	0.1	0.9360
S_{14}	0.9	0.1	0.0	0.0	0.6230
S_{15}	0.5	0.2	0.1	0.2	0.9100
S_{16}	0.5	0.3	0.0	0.2	0.8900
S_{17}	0.5	0.0	0.1	0.4	0.7205
S_{18}	0.7	0.2	0.1	0.0	0.7807
S_{19}	0.1	0.7	0.1	0.1	0.6217
S_{20}	0.3	0.6	0.1	0.0	0.7390
$\mathbf{S}_{c2}$	0.44	0.29	0.13	0.14	1.0000

Table 1: A list of given sources of evidences.

[9]Let's denote $k = |G^\Theta|$ the cardinality of G^Θ and consider S independent sources of evidence. If all sources are equireliable, the barycentre of belief masses of the S sources is given by: $\forall j = 1, \ldots, k$, $\bar{m}(X_j) = \frac{1}{S} \sum_{s=1}^{S} m_s(X_j)$, see [18] for details.

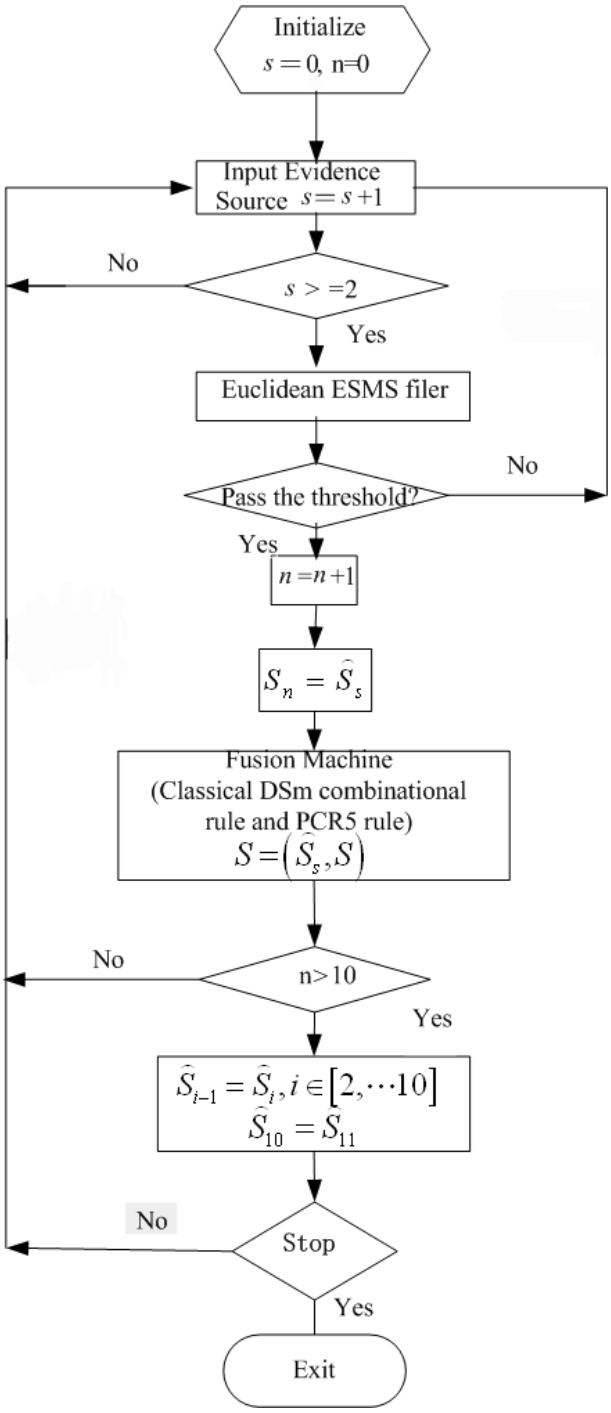


Figure 2: The procedure flow chart of DSmT-based generalized fusion machine.

In order to embed ESMS function to GFM [19], we give its working principle summarized in Figure 2, whose main steps of the algorithm for implementing the GFM are enumerated as follows:

1) Initialization of the parameters: the number of sources of evidence is set to zero, (i.e. one has initially no source, $s = 0$), so that the number of sources in the filter window is $n = 0$.
2) Include[10] a source of evidence $\hat{S}_s$ and then test if the number of sources s is less than 2. If $s \geq 2$, then go to next step, otherwise include/take into account another source of evidence $\hat{S}_2$.
3) Based on the barycentre of gbba of the front $n \leq 10$ evidence sources, the degrees of similarity are computed according to the formula (8), and compared with a prior tuned threshold. If it is larger than the threshold, then let $n = n + 1$. Otherwise, introduce a new source of evidence $\hat{S}_{s+1}$.
4) If $n = 1$, the current source, say S, is not involved in the fusion process. If $n = 2$, then the fusion step must apply between S and $\hat{S}_2$ with classical DSm rule [31], i.e. the conjunctive consensus. We then use PCR5 rule [32] to redistribute the remaining partial conflicts only to the sets involved in the corresponding partial conflicts. We get a new combined source affected with same index S. If $2 < n \leq 10$, after the current evidence source $\hat{S}_s$ is combined with the final source of evidence produced last time, a new source of evidence is obtained and assigned to S again. Whenever $n \leq 10$, go back to step 2), otherwise, the current source of evidence $\hat{S}_s$ under test has been accepted by the ESMS filter, $\hat{S}_i$ is assigned to $\hat{S}_{i-1}, i \in [2, s, 10]$, and $\hat{S}_s$ is assigned to $\hat{S}_{10}$. Then, $\hat{S}_{10}$ is combined with the last source S, the combined result is reassigned[11] to S, and then, go back to step 2).
5) Test whether to stop or not[12]: if no, then introduce a new source of evidence $\hat{S}_{s+1}$, otherwise stop and exit.

We show two simulation results in Figures 3 and 4 following the working principle of GFM, when we use the sources of evidences listed in Table 1.

[10]We assume the free DSm model and consider that the general basic belief assignments are given.

[11]In this work, we use also an ESMS filter window in a sliding mode.

[12]In our experiment, judge whether the mobile robot stops receiving sonar's data.

The comparison of Figures 3 and 4 yields the following remarks:

1) On the Figure 3, we don't see the real advantage of ESMS filter since the convergence to θ_1 without ESMS filter (green curve) is better than with ESMS filter in terms of improving fusion precision. This is because some useful information sources are filtered and thrown away with the increase of the threshold, the number of information sources to enter into the final fusion will become fewer and fewer. Generally speaking, this yields a slower speed of convergence to θ_1. For example, for one source S_3 in the Table 1, if S_3 is combined with itself once according to (6), the combinational result is $S = [m_N(\theta_1), m_N(\theta_2), m_N(\theta_1 \cap \theta_2), m_N(\theta_1 \cup \theta_2)]$[13] $= [0.5707, 0.0993, 0.1680, 0.1620]$. If twice, the result is $S = [0.6765, 0.0852, 0.1429, 0.0954]$. If thrice, then the result is $S = [0.7453, 0.00693, 0.1216, 0.0638]$. More the combinational times is, nearer by 1 $m_N(\theta_1)$ is and nearer by 0 $m_N(\theta_2)$ is .Therefore, ESMS filter might also result in losing some useful information, while it filter some bad information.
2) On the Figure 4, one sees the role played by the ESMS filter. When there are highly conflicting sources, the result of the fusion process will not converge if the ESMS filter is not used. With the fine tuning of the ESMS threshold, the convergence becomes better and better because ESMS filter processes the fused information, and withdraws the sources which might cause the results to be incorrect or imprecise, so that it improves the fusion precision and correctness.
3) The reduction of computing burden is obtained. Even if we have introduced an ESMS preprocessing step, it turns out that finally a drastic reduction of computing burden is obtained because we can significantly reduce the number of sources to combine with ESMS criterion.
4) We increase the applicability of the classical rules of combination. Since we reduce the number of conflicting sources of evidence thanks to the ESMS preprocessing, the degree of conflict between sources to combine is kept low. Therefore, classical rules, like Dempster's rule, which do not perform well in high conflicting situations can be used also in applications like the one studied here. Without ESMS preprocessing step, the classical fusion rules cannot work very well [23, 29]. Therefore, we extend their domain of applicability when using ESMS filtering step.

[13] $m_N(\cdot)$ refers to the new generalized basic belief assignment

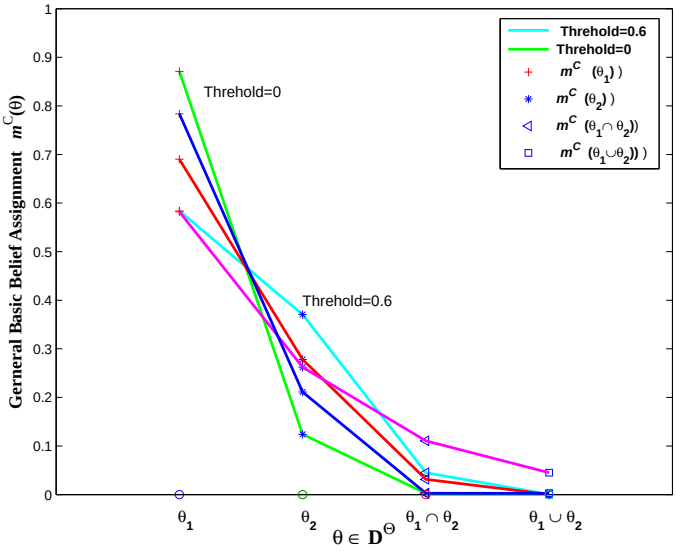


Figure 3: Fusion result of the front 10 sources using different thresholds ($0 \sim 0.9$).

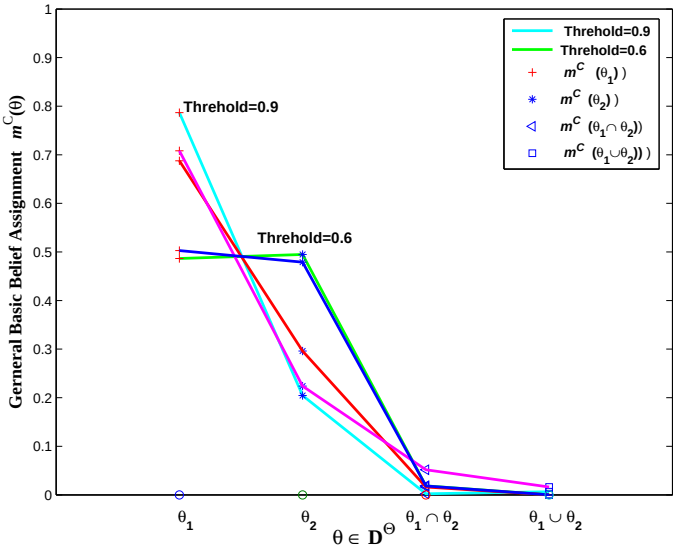


Figure 4: Fusion result of the total 20 sources using different thresholds ($0 \sim 0.9$).

7. An application in mobile robot perception

The information acquired in building grid map using sonar sensors on a mobile robot is usually uncertain, imprecise and even highly conflicting. Such application in autonomous robot perception and navigation provides a good platform to verify experimentally the capability of the ESMS filter in GFM. Although there exist many methods of building map based either on Probability theory [36], FST (Fuzzy System Theory) [24], DST [34], GST (Grey System Theory) [39, 40, 41], or DSmT [20], we just compare the performances of the map building using a classical fusion machine without ESMS filter (i.e. CFMW) with respect to the classical fusion machine with ESMS filter (called GFM) in the DSmT framework only. A detailed comparison between our current ESMS-based approach with other methods is given in a companion paper in [21] where we show that ESMS-based approach outperforms other approaches using almost the same experimental conditions and inputs. In order to further reduce the measurement noises, we improve our past belief assignment model of sonar sensors in DSmT framework[14] as follows:

$$m(\theta_1) = \begin{cases} (1-\rho/(R-2\epsilon)) \times (1-\lambda/2) & \text{if} \quad \begin{cases} R_{min} \leq \rho \leq R-\epsilon \\ 0 \leq \varphi \leq \omega/2 \end{cases} \\ 0 \qquad \text{otherwise} \end{cases} \tag{12}$$

$$m(\theta_2) = \begin{cases} \exp(-3(\rho-R)^2) \times \lambda & \text{if} \quad \begin{cases} R_{min} \leq \rho \leq R+\epsilon \\ 0 \leq \varphi \leq \omega/2 \end{cases} \\ 0 \qquad \text{otherwise} \end{cases} \tag{13}$$

[14] We assume that there are only two focal elements θ_1 and θ_2 in the frame of discernment. Elements of hyper-power set are θ_1,θ_2, $\theta_1 \cap \theta_2$ and $\theta_1 \cup \theta_2$. θ_1 represents the emptiness of a given grid cell, θ_2 represents the occupancy for a given grid cell, $\theta_1 \cap \theta_2$ means that there is some conflict between two sonar measurements for the same grid cell and $\theta_1 \cup \theta_2$ represents the ignorance for a grid cell because of the possible lack of measurement acquisition.

$$m(\theta_1 \cap \theta_2) = \begin{cases} 1 - (2(\rho - (R-\epsilon)/R)^2) & \text{if} \quad \begin{cases} R_{min} \leq \rho \leq R + \epsilon \\ 0 \leq \varphi \leq \omega/2 \end{cases} \\ 0 \qquad \text{otherwise} \end{cases} \tag{14}$$

$$m(\theta_1 \cup \theta_2) = \begin{cases} \tan(2(\rho - R)) \times (1 - \lambda) & \text{if} \quad R \leq \rho \leq R + \epsilon \\ 0 \qquad \text{otherwise} \end{cases} \tag{15}$$

Where, λ is given by (see [9] for justification)

$$\lambda = \begin{cases} 1 - (2\varphi/\omega)^2 & \text{if} \quad 0 \leq |\varphi| \leq \omega/2 \\ 0 \quad \text{otherwise} \end{cases} \tag{16}$$

The parameters R, ρ, ϵ, R_{min}, ω, and φ in formulas (12)-(16) were defined and used in [19, 20, 21]. R is the range measurement. ρ is the distance between the grid cell and sonar's emitting point. ϵ is the range measurement error. R_{min} is the minimal range distance of sonar sensors. ω is the scattering angle of sonar. φ is the angle between the line (from the grid cell to sonar emitting point) and the sonar's emitting direction. The following functions C_1 and C_2 play an important role in reducing noises in the process of map building. C_1 function, proposed by Wang in [39, 40, 41], is a constrict function[15] for sonar measurements defined by:

$$C_1 = \begin{cases} 0 & \text{if} \quad \rho > \rho_{l_2} \\ \frac{\rho_{l_2} - \rho}{\rho_{l_2} - \rho_{l_1}} & \text{if} \quad \rho_{l_1} \leq \rho \leq \rho_{l_2} \\ 1 & \text{if} \quad \rho < \rho_{l_2} \end{cases} \tag{17}$$

Where, ρ_{l_1} and ρ_{l_2} represents the upper and lower limits of valid measurements. C_2 is the constraint function[16] for sonar's uncertainty defined as follows:

$$C_2 = \begin{cases} (\frac{\rho - R + 0.5\epsilon}{0.5\epsilon})^2 & \text{if} \quad \rho - R > -0.5\epsilon \\ (\frac{\rho - R - 0.5\epsilon}{0.5\epsilon})^2 & \text{if} \quad \rho - R < 0.5\epsilon \\ 0 & \text{if} \quad \mid \rho - R \mid > 0.5\epsilon \end{cases} \tag{18}$$

[15]The main idea is that the sonar readings must be discounted according to sonar characteristics.

[16]The main idea consists in committing high belief assignments to sonar readings close to the sonar sensor.

Where, the product of C_1 and C_2 is multiplied by the belief assignment function, i.e. $m(\theta_2), m(\theta_1 \cap \theta_2), m(\theta_1 \cup \theta_2)$ respectively.

The experiment is performed by running Pioneer II mobile robot with 16 sonar detectors in the indoor laboratory environment as shown in Figure 5. The environment's size is $4550mm \times 3750mm$. The environment is divided into 91×75 rectangular cells having the same size according to the grid map method. The robot starts to move from the location (1 m, 0.6 m), which faces towards 0 degree. We take the corner of left bottom as the global coordinate origin of the map. Objects/obstacles in rectangular grid map are sketched in Figure 6. The processing steps of our intelligent perception and fusion system have been implemented with our software Toolbox developed under C++ 6.0 and with OpenGL server as a client end. When the robot moves in the environment, the server end collects much information (i.e. the location of robot, sensors measurements, velocity, etc.) from the mobile robot and its sensors onboard. Through the protocol of TCP/IP, the client end can get any information from the server end and fuse them.

Since our environment is small,the robot moves less time or a short distance. Then one only considers the self-localization method based on δ-NFAM[17] method [17, 19] with the search from $\theta - \delta_\theta$ to $\theta + \delta_\theta$. In order to reduce the computation burden, the restricted spreading arithmetic has been used. The flow chart of this procedure for this experiment is given in Figure 7. Its main steps are the following ones:

1) Initialize the parameters of the robot (location, velocity, etc.).
2) Acquire 16 sonar measurements, and robot's location from odometer, when the robot is running in the environment. We can calibrate the robot's pose with our δ-NFAM method [17, 19]. Here we set the first timer, of which interval is 100 ms.
3) Compute gbba of the fan-form area detected by each sonar sensor according to the formulas in [20].
4) Apply the DSmT-based GFM, that is, adopt Euclidean information filter to choose basic consistent sources of evidence according to the formula (8). Then combine the consistent sources with the DSm conjunctive rule [4, 5, 31] and compute gbbas after combination. Then,

[17] One has taken $\delta_\theta = 5^o$ in our experiment.

redistribute partial conflicting masses to the gbbas of sets involved in the partial conflict only with the PCR5 rule [32].

5) Compute the belief of occupancy $Bel(\theta_2)$ of some grid cells according to [31]. Save them into the map matrix and then go to step 6).
6) Update the map of the environment (here we set the second timer, of whose interval is 100 ms). Generally speaking, more the times of scanning map are, more accurate the final map reconstructed is. At the same time, also test whether the robot stops receiving the detecting data: if yes, then stop fusion and exit, otherwise, go back to step 2).

Figure 5: The real experimental environment.

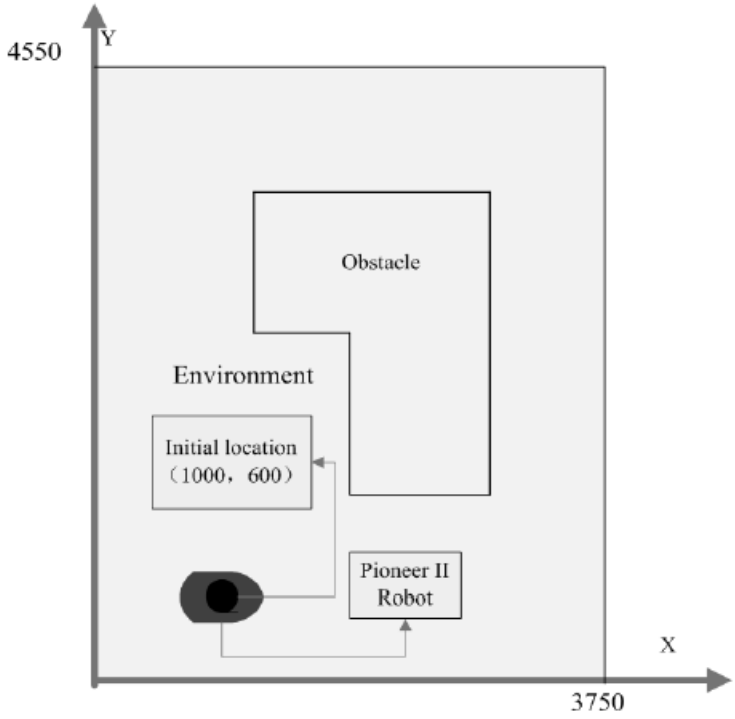


Figure 6: Global coordinate system for the experiment.

In this experiment, we obtain the maps built by the GFM before and after improving the sonar model as shown in Figure 8 and 9 respectively. In order to show the advantage of the ESMS filter in the GFM, we also compare our approach with the classical fusion machine which doesn't use the ESMS filter (called CFMW). The maps built by the CFMW before and after improving the sonar model are shown in Figures 10 and 11 respectively. Whenever the map is built before or after improving the sonar model, one sees that the GFM always outperforms the CFMW because one obtains clearer boundary outlines and less noises in the map reconstruction. In addition, the ESMS information filter coupled with PCR5 fusion rule, allows to reduce drastically the computational burden because ESMS filter can filter the outlier-sources. With the GFM approach, only the most consistent sources of evidence are combined and this allows to reduce the uncertainty in the fusion result and to improve the robot perception of the surrounded environment.

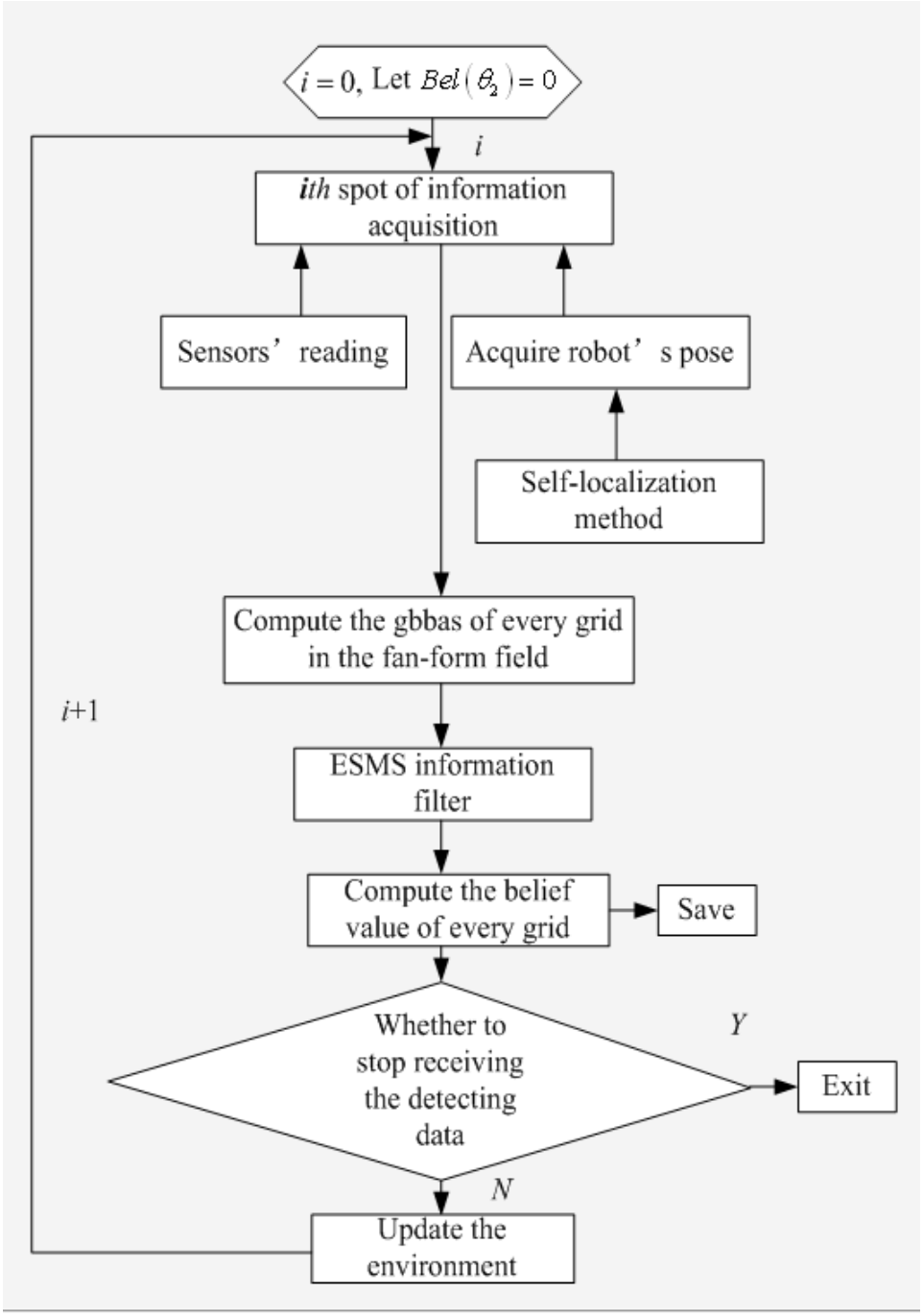


Figure 7: Flow chart of the map building with the GFM.

Figure 8: Map building based on the GFM before improving the sonar model.

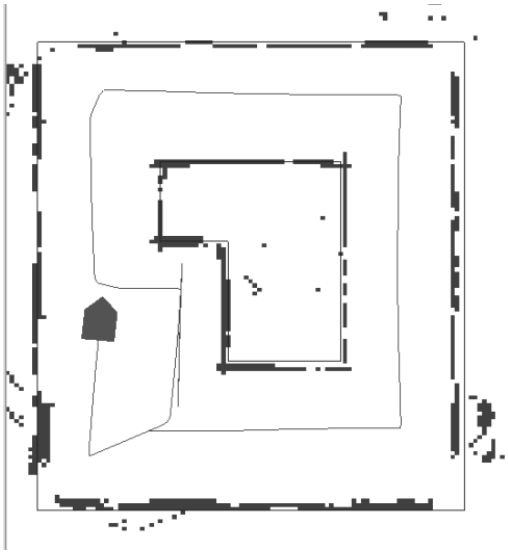


Figure 9: Map building based on the GFM after improving the sonar model.

Figure 10: Map building based on the CFMW before improving the sonar model.

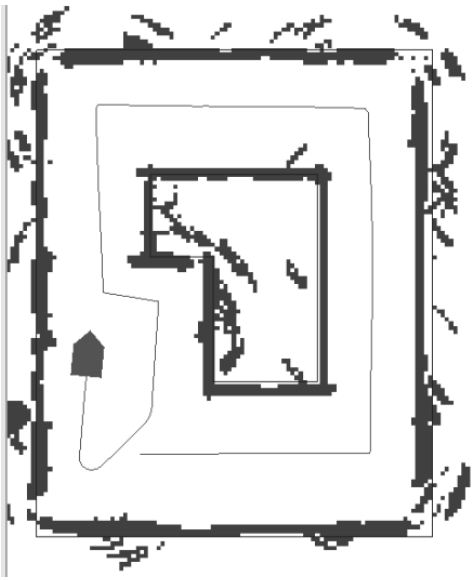


Figure 11: Map building based on the CFMW after improving the sonar model.

8. Conclusions

In this paper, a general Evidence Supporting Measure of Similarity (ESMS) between two basic belief assignments has been proposed. ESMS can be used on different fusion spaces (lattice structures) and with different distance measures. This approach allows to select the most coherent subset of sources of evidence available and to reject outlier-sources which are too conflicting with other sources. Hence, a drastic reduction of computational burden is possible with keeping good performances which is very attractive for real-time applications having limited computing resources. The hybrid of ESMS with the sophisticated and efficient PCR5 fusion rule of DSmT, called GFM (Generalized Fusion Machine), is specially useful and interesting in robotic applications involving real-time perception and navigation systems. The real application of GFM for mobile robot perception from sonar sensors presented in this work shows clearly a substantial improvement of the fusion result in map building/estimation of the surrounded environment. This work shows also the crucial role played by the most advanced fusion techniques for applications in robotics.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 60804063,60805032) and partially supported by The National High Technology Research and Development Program of China (863 Program, NO:2006AA040202)

References

[1] Bauer M., *Approximation algorithms and decision making in the Dempster-Shafer theory of evidence - an empirical study*, Int. J. of Approx. Reasoning. Vol. 17 (2-3), pp. 217, 1997.

[2] Chatzis V., Pitas I., *A generalized fuzzy mathematical morphology and its application in robust 2-D and 3-D object representation*, IEEE Transactions on Image Processing, pp.1798– 1810, 2000.

[3] Cover T. M., Thomas J. A., *Elements of Information Theory*, Wiley, New York, 1991.

[4] Dezert J., *Foundations for a new theory of plausible and paradoxical reasoning*, International Journal Information and Security, Vol. 9, pp.13–57, 2002.

[5] Dezert J., Smarandache F., *On the generation of hyper-power sets for the DSmT*, in Proc. of 6th International Conf. on Information Fusion, Cairns, Queensland, Australia, pp.1118–1125, 2003.

[6] Diaz J., Rifqi M., Bouchon-Meunier B., *A similarity measure between basic belief assignments*, in Proc. of International Conf. on Information fusion, Fusion 2006, Firenze, Italy, July 2006.

[7] Dilworth R.P., *Lattice theory*, Rhode Island American Methematical Society, 1961.

[8] Dubois D., Prade H., *Possibility theory, probability theory and multiple-valued logics: a clarification*, Annals of Mathematics and Artificial Intelligence, Vol. 32, pp. 35-66, 2001.

[9] Elfes A., Moravec H., *High resolution maps from wide angle sonar*, IEEE Int Conf. on Robotics and Automation, pp.116-121,1985.

[10] Grätzer G., *General lattice theory*, Academic Press, New York, 1978.

[11] ISIF (International Society of Information Fusion) web site: www.isif.org.

[12] Jeffrey R., *The logic of decision*, McGraw-Hill, New York, 1965.

[13] Jeffrey R., *Probability and the art of judgment*, Cambridge University Press, Cambridge, 1992.

[14] Jousselme A.-L., Grenier D., Bossé É., *A new distance between two bodies of evidence*, Information Fusion journal (Elsevier), Vol. 2, pp. 91–101, 2001.

[15] Kandasamy W.B.S., Smarandache F., *Basic Neutrosophic Algebraic Structures and Their Application to Fuzzy and Neutrosophic Models*, Hexis, Church Rock, 2004.

[16] Kolmogorov A. N., *Foundations of the theory of probability*, Chelsea Publishing Company, New York, 1960.

[17] Li X., *Research on Fusion Method of Imperfect Information from Multi-source and Its Application*, Ph. D. Thesis, Huazhong University of Science and Technology, China, June 2007.

[18] Li X., Dezert J., Huang X., *Selection of sources as a prerequisite for information fusion with application to SLAM*, in Proc. of the 9th International Conf. on Information fusion, Florence, Italy, July 10–13, 2006.

[19] Li X.,Huang X., Dezert J.,Wu Z., Zhang H., Xiong Y., *DSmT-based Generalized Fusion Machine for information fusion in robot map building*, in Proc. of Int. Colloq. on Information Fusion 2007 (ICIF' 2007), pp. 63 –70, Xi'an, China, August 22-25, 2007.

[20] Li X.,,Huang X., Dezert J.,Duan L., Wang M., *A successful application of DSmT in sonar grid map building and comparison with DST-based approach*, Int. J. of Innovative Comp., Inf. and Control, Vol. 3, No. 3, pp. 539–551, 2007.

[21] Li X., Zhu B., Dezert J., Dai X., *An improved fusion machine for robot perception*, to be accepted by Journal of intelligent and robotic systems after revision.

[22] Loschi R.H., Iglesias P.L., *Jeffrey's rule: An alternative procedure to model uncertainty*, Estadística, Vol. 57(168,169), pp. 11-26, 2005.

[23] Murphy C. K, *Combining belief functions when evidence con flicts*. Decision Support Systems ,29(1):1-9, 2000 .

[24] Ofir C., Yael E., *Adaptive fuzzy logic algorithm for grid-map based sensor fusion*, IEEE Intell. Vehicles Symp., Parma, Italy, pp. 625–630, June 2004.

[25] Papoulis A., *Probability, Random Variables and Stochastic Processes*, McGraw Hill, 4th Revised Edition, 2002.

[26] Petit-Renaud S., *Application de la théorie des croyances et des systèmes flous à l'estimation fonctionnelle en présence d'informations incertaines ou imprécises*, Ph.D. Thesis, Université de Technologie de Compiègne, France, December 1999.

[27] Ristic B., Smets Ph., *Association of Uncertain Combat ID Declarations*, in Proc. of Cogis '06 Conference, Paris, France, March 2006.

[28] Ristic B., Smets Ph., *The TBM global distance measure for the association of uncertain combat IDdeclarations*, in Proc. of International Information Fusion Conference, Fusion 2006, Florence, Italy, July 2006.

[29] Shafer G., *A mathematical theory of evidence*, Princeton University Press, Princeton, NJ, 1976.

[30] Smarandache F., *Unification of Fusion Theories (UFT)*, International Journal of Applied Mathematics and Statistics, Vol.2, pp. 1–14, Dec. 2004.

[31] Smarandache F., Dezert J. (Editors), *Advances and Applications of DSmT for Information Fusion (Collected works)*, Vol.1, American Research Press, Rehoboth, June 2004.
http://www.gallup.unm.edu/~smarandache/DSmT-book1.pdf.

[32] Smarandache F., Dezert J. (Editors), *Advances and Applications of DSmT for Information Fusion (Collected works)*, Vol.2, American Research Press, Rehoboth, 2006.
http://www.gallup.unm.edu/~smarandache/DSmT-book2.pdf.

[33] Smarandache F., Dezert J. (Editors), *Advances and Applications of DSmT for Information Fusion (Collected works)*, Vol.3, American Research Press, Rehoboth, 2009.
http://www.gallup.unm.edu/~smarandache/DSmT-book3.pdf.

[34] Su L., Cao Z., Wang S., Tan M., *A real-time on-line method for exploring unknown environment with multiple robots*, High Technology Letters, 11, pp. 56–60, 2003.

[35] Tessem B., *Approximations for effecent computation in the theory of evidence*, Artificial Intelligence, Vol. 61, pp. 315–329, 1993.

[36] Thrun S., Fox D., Burgard W., *A probabilistic approach to concurrent mapping and localization for mobile robots*, Machine Learning, Vol. 31(1-3), pp. 29–53, 1998.

[37] Tversky A., *Features of similarity*, Psychological Review, Vol. 84, pp. 327–352, 1977.

[38] Walley P., *Statistical reasoning with imprecise probabilities*, Chapman and Hall, London, 1991.

[39] Wang W., Chen W., Xi Y., *Uncertainty sensor information processing in map building of mobile robot*, Acta Automatica Sinica, Vol. 29(2), pp. 267–274, 2003.

[40] Wang W., Xi Y., Chen W., *Map building for a mobile robot based on grey system theory*, J. of Systems Eng. and Elec., Vol. 14, No. 3, pp. 67–72, 2003.

[41] Wang W., *Map building by mobile robots in unknown environment*, Ph. D. Thesis, Shanghai Jiao Tong University, China, 2003.

[42] Zadeh L., *Fuzzy sets as the basis for a theory of possibility*, Fuzzy Sets and Systems, Vol. 1, pp. 3–28, 1978 (Reprinted in Fuzzy Sets and Systems 100 (Supplement) pp. 9–34, 1999.)

[43] Zouhal L.M., Denœux T., *An evidence-theoric k-NN rule with parameter optimization*, IEEE trans. Syst. Man Cybern. C, Vol. 28, No. 2, pp. 263–271, 1998.

Contradiction Measures and Specificity Degrees of Basic Belief Assignments

Florentin Smarandache
Arnaud Martin
Christophe Osswald

Originally published as: Smarandache F., Martin A., Osswald C - *Contradiction Measures and Specificity Degrees of Basic Belief Assignments,* in Proceedings of the 14th International Conference on Information Fusion, Chicago, IL, USA, 5-8 July 2011, and reprinted with permission.

***Abstract*—In the theory of belief functions, many measures of uncertainty have been introduced. However, it is not always easy to understand what these measures really try to represent. In this paper, we re-interpret some measures of uncertainty in the theory of belief functions. We present some interests and drawbacks of the existing measures. On these observations, we introduce a measure of contradiction. Therefore, we present some degrees of non-specificity and Bayesianity of a mass. We propose a degree of specificity based on the distance between a mass and its most specific associated mass. We also show how to use the degree of specificity to measure the specificity of a fusion rule. Illustrations on simple examples are given.**

Keywords: Belief function, uncertainty measures, specificity, conflict.

I. Introduction

The theory of belief functions was first introduced by [1] in order to represent some imprecise probabilities with *upper* and *lower probabilities*. Then [15] proposed a mathematical theory of evidence.

Let Θ be a frame of discernment. A *basic belief assignment* (bba) m is the mapping from elements of the powerset 2^Θ onto $[0,1]$ such that:

$$\sum_{X\in 2^\Theta} m(X) = 1. \quad (1)$$

The axiom $m(\emptyset) = 0$ is often used, but not mandatory. A *focal element* X is an element of 2^Θ such that $m(X) \neq 0$.. The difference of a bba with a probability is the domain of definition. A bba is defined on the powerset 2^Θ and not only on Θ. In the powerset, each element is not equivalent in terms of precision. Indeed, $\theta_1 \in \Theta$ is more precise than $\theta_1 \cup \theta_2 \in 2^\Theta$.

In the case of the DSmT introduced in [17], the bba are defined on an extension of the powerset: the hyper powerset noted D^Θ, formed by the closure of Θ by union and intersection. The problem of signification of each focal element is the same as in 2^Θ. For instance, $\theta_1 \in \Theta$ is less precise than $\theta_1 \cap \theta_2 \in D^\Theta$. In the rest of the paper, we will note G^Θ for either 2^Θ or D^Θ.

In order to try to quantify the measure of uncertainty such as in the set theory [5] or in the theory of probabilities [16], some measures have been proposed and discussed in the theory of belief functions [2], [7], [8], [21]. However, the domain of definition of the bba does not allow an ideal definition of measure of uncertainty. Moreover, behind the term of uncertainty, different notions are hidden.

In the section II, we present different kinds of measures of uncertainty given in the state of art, we discuss them and give our definitions of some terms concerning the uncertainty. In section III, we introduce a measure of contradiction and discuss it. We introduce simple degrees of uncertainty in the section IV, and propose a degree of specificity in the section V. We show how this degree of specificity can be used to measure the specificity of a combination rule.

II. Measures of uncertainty on belief functions

In the framework of the belief functions, several functions (we call them *belief functions*) are in one to one correspondence with the bba: bel, pl and q. From these belief functions, we can define several measures of uncertainty. Klir in [8] distinguishes two kinds of uncertainty: the non-specificity and the discord. Hence, we recall hereafter the main belief functions, and some non-specificity and discord measures.

A. Belief functions

Hence, the credibility and plausibility functions represent respectively a minimal and maximal belief. The *credibility* function is given from a bba for all $X \in G^\Theta$ by:

$$\mathrm{bel}(X) = \sum_{Y\subseteq X, Y\neq\emptyset} m(Y). \quad (2)$$

The *plausibility* is given from a bba for all $X \in G^\Theta$ by:

$$\mathrm{pl}(X) = \sum_{Y\in G^\Theta, Y\cap X\neq\emptyset} m(Y). \quad (3)$$

The *commonality* function is also another belief function given by:

$$\mathrm{q}(X) = \sum_{Y\in G^\Theta, Y\supseteq X} m(Y). \quad (4)$$

These functions allow an implicit model of imprecise and uncertain data. However, these functions are monotonic by inclusion: bel and pl are increasing, and q is decreasing. This

is the reason why the most of time we use a probability to take a decision. The most used projection into probability subspace is the pignistic probability transformation introduced by [18] and given by:

$$\mathrm{betP}(X) = \sum_{Y \in G^\Theta, Y \neq \emptyset} \frac{|X \cap Y|}{|Y|} m(Y), \tag{5}$$

where $|X|$ is the cardinality of X, in the case of the DSmT that is the number of disjoint elements corresponding in the Venn diagram.

From this probability, we can use the measure of uncertainty given in the theory of probabilities such as the Shannon entropy [16], but we loose the interest of the belief functions and the information given on the subsets of the discernment space Θ.

B. Non-specificity

The non-specificity in the classical set theory is the imprecision of the sets. Such as in [14], we define in the theory of belief functions, the non-specificity related to vagueness and non-specificity.

Definition *The* non-specificity *in the theory of belief functions quantifies how a bba m is imprecise.*

The non-specificity of a subset X is defined by Hartley [5] by $\log_2(|X|)$. This measure was generalized by [2] in the theory of belief functions by:

$$\mathrm{NS}(m) = \sum_{X \in G^\Theta,\ X \neq \emptyset} m(X) \log_2(|X|). \tag{6}$$

That is a weighted sum of the non-specificity, and the weights are given by the basic belief in X. Ramer in [13] has shown that it is the unique possible measure of non-specificity in the theory of belief functions under some assumptions such as symmetry, additivity, sub-additivity, continuity, branching and normalization.

If the measure of the non-specificity on a bba is low, we can consider the bba is specific. Yager in [21] defined a specificity measure such as:

$$S(m) = \sum_{X \in G^\Theta,\ X \neq \emptyset} \frac{m(X)}{|X|}. \tag{7}$$

Both definitions corresponded to an accumulation of a function of the basic belief assignment on the focal elements. Unlike the classical set theory, we must take into account the bba in order to quantify (to weight) the belief of the imprecise focal elements. The imprecision of a focal element can of course be given by the cardinality of the element.

First of all, we must be able to compare the non-specificity (or specificity) between several bba's, event if these bba's are not defined on the same discernment space. That is not the case with the equations (6) and (7). The non-specificity of the equation (6) takes its values in $[0, \log_2(|\Theta|)]$. The specificity of the equation (7) can have values in $[\frac{1}{|\Theta|}, 1]$. We will show how we can easily define a degree of non-specificity in $[0, 1]$. We could also define a degree of specificity from the equation (7), but that is more complicated and we will later show how we can define a specificity degree.

The most non-specific bba's for both equations (6) and (7) are the total ignorance bba given by the categorical bba $m_\Theta : m(\Theta) = 1$. We have $\mathrm{NS}(m) = \log_2(|\Theta|)$ and $S(m) = \frac{1}{|\Theta|}$. This categorical bba is clearly the most non-specific for us. However, the most specific bba's are the Bayesian bba's. The only focal elements of a Bayesian bba are the simple elements of Θ. On these kinds of bba m we have $\mathrm{NS}(m) = 0$ and $S(m) = 1$. For example, we take the three Bayesian bba's defined on $\Theta = \{\theta_1, \theta_2, \theta_3\}$ by:

$$m_1(\theta_1) = m_1(\theta_2) = m_1(\theta_3) = 1/3, \tag{8}$$

$$m_2(\theta_1) = m_2(\theta_2) = 1/2,\ m_2(\theta_3) = 0, \tag{9}$$

$$m_3(\theta_1) = 1,\ m_3(\theta_2) = m_3(\theta_3) = 0. \tag{10}$$

We obtain the same non-specificity and specificity for these three bba's.

That hurts our intuition; indeed, we intuitively expect that the bba m_3 is the most specific and the m_1 is the less specific. We will define a degree of specificity according to a most specific bba that we will introduce.

C. Discord

Different kinds of discord have been defined as extensions for belief functions of the Shannon entropy, given for the probabilities. Some discord measures have been proposed from plausibility, credibility or pignistic probability:

$$E(m) = -\sum_{X \in G^\Theta} m(X) \log_2(\mathrm{pl}(X)), \tag{11}$$

$$C(m) = -\sum_{X \in G^\Theta} m(X) \log_2(\mathrm{bel}(X)), \tag{12}$$

$$D(m) = -\sum_{X \in G^\Theta} m(X) \log_2(\mathrm{betP}(X)), \tag{13}$$

with $E(m) \leq D(m) \leq C(m)$. We can also give the Shanon entropy on the pignistic probability:

$$-\sum_{X \in G^\Theta} \mathrm{betP}(X) \log_2(\mathrm{betP}(X)). \tag{14}$$

Other measures have been proposed, [8] has shown that these measures can be resumed by:

$$-\sum_{X \in G^\Theta} m(X) \log_2(1 - \mathrm{Con}_m(X)), \tag{15}$$

where Con is called a conflict measure of the bba m on X. However, in our point of view that is not a conflict such presented in [20], but a contradiction. We give the both following definitions:

Definition *A* contradiction *in the theory of belief functions quantifies how a bba m contradicts itself.*

Definition (C1) *The* conflict *in the theory of belief functions can be defined by the contradiction between 2 or more bba's.*

In order to measure the conflict in the theory of belief functions, it was usual to use the mass k given by the

conjunctive combination rule on the empty set. This rule is given by two basic belief assignments m_1 and m_2 and for all $X \in G^\Theta$ by:

$$m_{\text{c}}(X) = \sum_{A \cap B = X} m_1(A) m_2(B) := (m_1 \oplus m_2)(X). \quad (16)$$

$k = m_{\text{c}}(\emptyset)$ can also be interpreted as a non-expected solution.

In [21], Yager proposed another conflict measure from the value of k given by $-\log_2(1-k)$.

However, as observed in [9], the weight of conflict given by k (and all the functions of k) is not a conflict measure between the basic belief assignments. Indeed this value is completely dependant of the conjunctive rule and this rule is non-idempotent - the combination of identical basic belief assignments leads generally to a positive value of k. To highlight this behavior, we defined in [12] the *auto-conflict* which quantifies the intrinsic conflict of a bba. The auto-conflict of order n for one expert is given by:

$$a_n = \left(\bigoplus_{i=1}^{n} m \right)(\emptyset). \quad (17)$$

The auto-conflict is a kind of measure of the contradiction, but depends on the order. We studied its behavior in [11]. Therefore we need to define a measure of contradiction independent on the order. This measure is presented in the next section III.

III. A CONTRADICTION MEASURE

The definition of the conflict (C1) involves firstly to measure it on the bba's space and secondly that if the opinions of two experts are far from each other, we consider that they are in conflict. That suggests a notion of distance. That is the reason why in [11], we give a definition of the measure of conflict between experts assertions through a distance between their respective bba's. The conflict measure between 2 experts is defined by:

$$\text{Conf}(1,2) = d(m_1, m_2). \quad (18)$$

We defined the conflict measure between one expert i and the other $M-1$ experts by:

$$\text{Conf}(i, \mathcal{E}) = \frac{1}{M-1} \sum_{j=1, i \neq j}^{M} \text{Conf}(i,j), \quad (19)$$

where $\mathcal{E} = \{1, \ldots, M\}$ is the set of experts in conflict with i. Another definition is given by:

$$\text{Conf}(i, M) = d(m_i, \overline{m_M}), \quad (20)$$

where $\overline{m_M}$ is the bba of the artificial expert representing the combined opinions of all the experts in $\mathcal{E}$ except i.

We use the distance defined in [6], which is for us the most appropriate, but other distances are possible. See [4] for a comparison of distances in the theory of belief functions. This distance is defined for two basic belief assignments $\mathbf{m}_1$ and $\mathbf{m}_2$ on G^Θ by:

$$d(m_1, m_2) = \sqrt{\frac{1}{2}(\mathbf{m}_1 - \mathbf{m}_2)^T \underline{\underline{D}} (\mathbf{m}_1 - \mathbf{m}_2)}, \quad (21)$$

where $\underline{\underline{D}}$ is an $G^{|\Theta|} \times G^{|\Theta|}$ matrix based on Jaccard distance whose elements are:

$$D(A,B) = \begin{cases} 1, \text{ if } A = B = \emptyset, \\ \dfrac{|A \cap B|}{|A \cup B|}, \ \forall A, B \in G^\Theta. \end{cases} \quad (22)$$

However, this measure is a *total conflict* measure. In order to define a contradiction measure we keep the same spirit. First, the contradiction of an element X with respect to a bba m is defined as the distance between the bba's m and m_X, where $m_X(X) = 1$, $X \in G^\Theta$, is the categorical bba:

$$\text{Contr}_m(X) = d(m, m_X), \quad (23)$$

where the distance can also be the Jousselme distance on the bba's. The contradiction of a bba is then defined as a weighted contradiction of all the elements X of the considered space G^Θ:

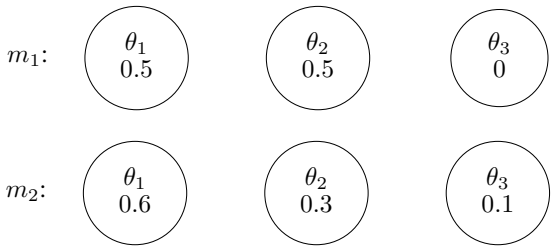
$$\text{Contr}_m = 2 \sum_{X \in G^\Theta} m(X) d(m, m_X). \quad (24)$$

The factor 2 is given to obtain values in $[0,1]$ as shown in the following illustration.

A. Illustration

First we note that for all categorical bbas m_Y, the contradiction given by the equation (23) gives $\text{Contr}_{m_Y}(Y) = 0$ and the contradiction given by the equation (24) brings also $\text{Contr}_{m_Y} = 0$. Considering the bba $m_1(\theta_1) = 0.5$ and $m_1(\theta_2) = 0.5$, we have $\text{Contr}_{m_1} = 1$. That is the maximum of the contradiction, hence the contraction of a bba takes its values in $[0,1]$.

Figure 1. Bayesian bba's



Taking the Bayesian bba given by: $m_2(\theta_1) = 0.6$, $m_2(\theta_2) = 0.3$, and $m_2(\theta_3) = 0.1$. We obtain:

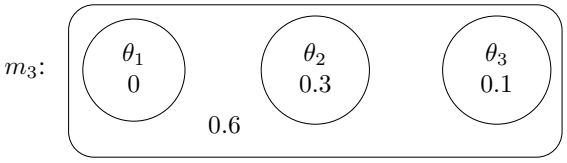
$$\begin{aligned} \text{Contr}_{m_2}(\theta_1) &\simeq 0.36, \\ \text{Contr}_{m_2}(\theta_2) &\simeq 0.66, \\ \text{Contr}_{m_2}(\theta_3) &\simeq 0.79 \end{aligned}$$

The contradiction for m_2 is $\text{Contr}_{m_2} = 0.9849$.

Take $m_3(\{\theta_1, \theta_2, \theta_3\}) = 0.6$, $m_3(\theta_2) = 0.3$, and $m_3(\theta_3) = 0.1$; the masses are the same than m_2, but the highest one is on $\Theta = \{\theta_1, \theta_2, \theta_3\}$ instead of θ_1. We obtain:

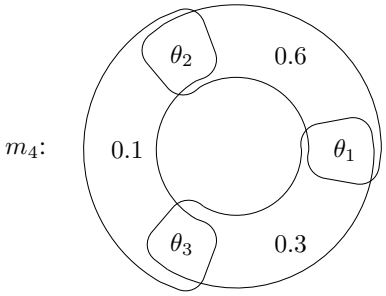
$$\begin{aligned} \text{Contr}_{m_3}(\{\theta_1, \theta_2, \theta_3\}) &\simeq 0.28, \\ \text{Contr}_{m_3}(\theta_2) &\simeq 0.56, \\ \text{Contr}_{m_3}(\theta_3) &\simeq 0.71 \end{aligned}$$

Figure 2. Non-dogmatic bba



The contradiction for m_3 is $\mathrm{Contr}_{m_3} = 0.8092$. We can see that the contradiction is lowest thanks to the distance taking into account the imprecision of Θ.

Figure 3. Focal elements of cardinality 2



If we consider now the same mass values but on focal elements of cardinality 2: $m_4(\{\theta_1, \theta_2\}) = 0.6$, $m_4(\theta_1, \theta_3) = 0.3$, and $m_4(\theta_2, \theta_3) = 0.1$. We obtain:

$$\begin{aligned} \mathrm{Contr}_{m_4}(\{\theta_1, \theta_2\}) &\simeq 0.29, \\ \mathrm{Contr}_{m_4}(\{\theta_1, \theta_3\}) &\simeq 0.53, \\ \mathrm{Contr}_{m_4}(\{\theta_2, \theta_3\}) &\simeq 0.65 \end{aligned}$$

The contradiction for m_4 is $\mathrm{Contr}_{m_4} = 0.80$.

Fewer of focal elements there are, smaller the contradiction of the bba will be, and more the focal elements are precise, higher the contradiction of the bba will be.

IV. Degrees of uncertainty

We have seen in the section II that the measure non-specificity given by the equation (6) take its values in a space dependent on the size of the discernment space Θ. Indeed, the measure of non-specificity takes its values in $[0, \log_2(|\Theta|)]$.

In order to compare some non-specificity measures in an absolute space, we define a degree of non-specificity from the equation (6) by:

$$\begin{aligned} \delta_{\mathrm{NS}}(m) &= \sum_{X \in G^\Theta,\ X \neq \emptyset} m(X) \frac{\log_2(|X|)}{\log_2(|\Theta|)} \\ &= \sum_{X \in G^\Theta,\ X \neq \emptyset} m(X) \log_{|\Theta|}(|X|). \end{aligned} \tag{25}$$

Therefore, this degree takes its values into $[0, 1]$ for all bba's m, independently of the size of discernment. We still have $\delta_{\mathrm{NS}}(m_\Theta) = 1$, where m_Θ is the categorical bba giving the total ignorance. Moreover, we obtain $\delta_{\mathrm{NS}}(m) = 0$ for all Bayesian bba's.

Table I
Evaluation of Bayesianity on examples

	m_1	m_2	m_3	m_4	m_5	m_6	m_Θ
θ_1	0.4	0.3	0.1	0.3	0	0	0
θ_2	0.1	0.1	0.3	0.1	0	0	0
θ_3	0.1	0.1	0.1	0.1	0	0	0
$\theta_1 \cup \theta_2$	0.3	0.3	0.5	0	0.6	0.6	0
$\theta_1 \cup \theta_3$	0.1	0.2	0	0	0.4	0	0
$\theta_2 \cup \theta_3$	0	0	0	0	0	0	0
Θ	0	0	0	0.5	0	0.4	1
δ_B	0.75	0.68	0.68	0.5	0.37	0.23	0
δ_{NS}	0.25	0.32	0.32	0.5	0.63	0.77	1

From the definition of the degree of non-specificity, we can propose a degree of specificity such as:

$$\begin{aligned} \delta_B(m) &= 1 - \sum_{X \in G^\Theta,\ X \neq \emptyset} m(X) \frac{\log_2(|X|)}{\log_2(|\Theta|)} \\ &= 1 - \sum_{X \in G^\Theta,\ X \neq \emptyset} m(X) \log_{|\Theta|}(|X|). \end{aligned} \tag{26}$$

As we observe already the degree of non-specificity given by the equation (26) does not really measure the specificity but the Bayesianity of the considered bba. This degree is equal to 1 for Bayesian bba's and not one for other bba's.

Definition *The* Bayesianity *in the theory of belief functions quantify how far a bba m is from a probability.*

We illustrate these degrees in the next subsection.

A. Illustration

In order to illustrate and discuss the previous introduced degrees we take some examples given in the table I. The bba's are defined on 2^Θ where $\Theta = \{\theta_1, \theta_2, \theta_3\}$. We only consider here non-Bayesian bba's, else the values are still given hereinbefore.

We can observe for a given sum of basic belief on the singletons of Θ the Bayesianity degree can change according to the basic belief on the other focal elements. For example, the degrees are the same for m_2 and m_3, but different for m_4. For the bba m_4, note that the sum of the basic beliefs on the singletons is equal to the basic belief on the ignorance. In this case the Bayesianity degree is exactly 0.5. That is conform to the intuitive signification of the Bayesianity. If we look m_5 and m_6, we first note that there is no basic belief on the singletons. As a consequence, the Bayesianity is weaker. Moreover, the bba m_5 is naturally more Bayesian than m_6 because of the basic belief on Θ.

We must add that these degrees are dependent on the cardinality of the frame of discernment for non Bayesian bba's. Indeed, if we consider the given bba m_1 with $\Theta = \{\theta_1, \theta_2, \theta_3\}$, the degree $\delta_B = 0.75$. Now if we consider the same bba

with $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ (no beliefs are given on θ_4), the Bayesianity degree is $\delta_B = 0.80$. The larger is the frame, the larger becomes the Bayesianity degree.

V. Degree of specificity

In the previous section, we showed the importance to consider a degree instead of a measure. Moreover, the measures of specificity and non-specificity given by the equations (7) and (6) give the same values for every Bayesian bba's as seen on the examples of the section II.

We introduce here a degree of specificity based on comparison with the bba the most specific. This degree of specificity is given by:

$$\delta_S(m) = 1 - d(m, m_s), \tag{27}$$

where m_s is the bba the most specific associated to m and d is a distance defined onto $[0, 1]$. Here we also choose the Jousselme distance, the most appropriated on the bba's space, with values onto $[0, 1]$. This distance is dependent on the size of the space G^Θ, we have to compare the degrees of specificity for bba's defined from the same space. Accordingly, the main problem is to define the bba the most specific associated to m.

A. The most specific bba

In the theory of belief functions, several partial orders have been proposed in order to compare the bba's [3]. These partial ordering are generally based on the comparisons of their plausibilities or their communalities, specially in order to find the least-committed bba. However, comparing bba's with plausibilities or communality can be complex and without unique solution.

The problem to find the most specific bba associated to a bba m does not need to use a partial ordering. We limit the specific bba's to the categorical bba's: $m_X(X) = 1$ where $X \in G^\Theta$ and we will use the following axiom and proposition:

Axiom *For two categorical bba's m_X and m_Y, m_X is more specific than m_Y if and only if $|X| < |Y|$.*

In case of equality, the both bba's are *isospecific*.

Proposition *If we consider two isospecific bba's m_X and m_Y, the Jousselme distance between every bba m and m_X is equal to the Jousselme distance between m and m_Y:*

$$\forall m,\ d(m, m_X) = d(m, m_Y) \tag{28}$$

if and only if $m(X) = m(Y)$.

Proof *The proof is obvious considering the equations* (21) *and* (22)*. As the both bba's m_X and m_Y are categoric there is only one non null term in the difference of vectors $m - m_X$ and $m - m_Y$. These both terms a_X and a_Y are equal, because m_X and m_Y are isospecific and so according to the equation* (22) *$D(Z, X) = D(Z, Y)\ \forall Z \in G^\Theta$. Therefore $m(X) = m(Y)$, that proves the proposition* □

We define *the most specific bba* m_s associated to a bba m as a categorical bba as follows: $m_s(X_{\max}) = 1$ where $X_{\max} \in G^\Theta$.

Therefore, the matter is now how to find $X_{\max}$. We propose two approaches:

First approach, Bayesian case

If m is a Bayesian bba we define $X_{\max}$ such as:

$$X_{\max} = \arg\max(m(X),\ X \in \Theta). \tag{29}$$

If there exist many $X_{\max}$ (*i.e.* having the same maximal bba and giving many isospecific bba's), we can take any of them. Indeed, according to the previous proposition, the degree of specificity of m will be the same with m_s given by either $X_{\max}$ satisfying (29).

First approach, non-Bayesian case

If m is a non-Bayesian bba, we can define $X_{\max}$ in a similar way such as:

$$X_{\max} = \arg\max\left(\frac{m(X)}{|X|},\ X \in G^\Theta,\ X \neq \emptyset\right). \tag{30}$$

In fact, this equation generalizes the equation (29). However, in this case we can also have several $X_{\max}$ not giving isospecific bba's. Therefore, we choose one of the more specific bba, *i.e.* believing in the element with the smallest cardinality. Note also that we keep the terms of Yager from the equation (7).

Second approach

Another way in the case of non-Bayesian bba m is to transform m into a Bayesian bba, thanks to one of the probability transformation such as the pignistic probability. Afterwards, we can apply the previous Bayesian case. With this approach, the most specific bba associated to a bba m is always a categorical bba such as: $m_s(X_{\max}) = 1$ where $X_{\max} \in \Theta$ and not in G^Θ.

B. Illustration

In order to illustrate this degree of specificity we give some examples. The table II gives the degree of specificity for some Bayesian bba's. The smallest degree of specificity of a Bayesian bba is obtained for the uniform distribution (m_1), and the largest degree of specificity is of course obtain for categorical bba (m_8).

The degree of specificity increases when the differences between the mass of the largest singleton and the masses of other singletons are getting bigger: $\delta_S(m_3) < \delta_S(m_4) < \delta_S(m_5) < \delta_S(m_6)$. In the case when one has three disjoint singletons and the largest mass of one of them is 0.45 (on θ_1), the minimum degree of specificity is reached when the masses of θ_2 and θ_3 are getting further from the mass of θ_1 (m_6). If two singletons have the same maximal mass, bigger this mass is and bigger is the degree of specificity: $\delta_S(m_2) < \delta_S(m_3)$.

In the case of non-Bayesian bba, we first take a simple example:

$$m_1(\theta_1) = 0.6, \quad m_1(\theta_1 \cup \theta_2) = 0.4 \tag{31}$$

$$m_2(\theta_1) = 0.5, \quad m_2(\theta_1 \cup \theta_2) = 0.5. \tag{32}$$

For these two bba's m_1 and m_2, both proposed approaches give the same most specific bba: m_{θ_1}. We obtain $\delta_S(m_1) =$

Table II
ILLUSTRATION OF THE DEGREE OF SPECIFICITY ON BAYESIAN BBA.

	θ_1	θ_2	θ_3	δ_S
m_1	1/3	1/3	1/3	0.423
m_2	0.4	0.4	0.2	0.471
m_3	0.45	0.45	0.10	0.493
m_4	0.45	0.40	0.15	0.508
m_5	0.45	0.3	0.25	0.523
m_6	0.45	0.275	0.275	0.524
m_7	0.6	0.3	0.1	0.639
m_8	1	0	0	1

0.7172 and $\delta_S(m_2) = 0.6465$. Therefore, m_1 is more specific than m_2. Remark that these degrees are the same if we consider the bba's defined on 2^{Θ_2} and 2^{Θ_3}, with $\Theta_2 = \{\theta_1, \theta_2\}$ and $\Theta_3 = \{\theta_1, \theta_2, \theta_3\}$. If we now consider Bayesian bba $m_3(\theta_1) = m_3(\theta_2) = 0.5$, the associated degree of specificity is $\delta_S(m_3) = 0.5$. As expected by intuition, m_2 is more specific than m_3.

We consider now the following bba:

$$m_4(\theta_1) = 0.6, \quad m_1(\theta_1 \cup \theta_2 \cup \theta_3) = 0.4. \tag{33}$$

The most specific bba is also m_{θ_1}, and we have $\delta_S(m_4) = 0.6734$. This degree of specificity is naturally smaller than $\delta_S(m_1)$ because of the mass 0.4 on a more imprecise focal element.

Let's now consider the following example:

$$m_5(\theta_1 \cup \theta_2) = 0.7, \quad m_5(\theta_1 \cup \theta_3) = 0.3. \tag{34}$$

We do not obtain the same most specific bba with both proposed approaches: the first one will give the categorical bba $m_{\theta_1 \cup \theta_2}$ and the second one m_{θ_1}. Hence, the first degree of specificity is $\delta_S(m_5) = 0.755$ and the second one is $\delta_S(m_5) = 0.111$. We note that the second approach produces naturally some smaller degrees of specificity.

C. Application to measure the specificity of a combination rule

We propose in this section to use the proposed degree of specificity in order to measure the quality of the result of a combination rule in the theory of belief functions. Indeed, many combination rules have been developed to merge the bba's [10], [19]. The choice of one of them is not always obvious. For a special application, we can compare the produced results of several rules, or try to choose according to the special proprieties wanted for an application. We also proposed to study the comportment of the rules on generated bba's [12]. However, no real measures have been used to evaluate the combination rules. Hereafter, we only show how we can use the degree of specificity to evaluate and compare the combination rules in the theory of belief functions. A complete study could then be done for example on generated bba's. We recall here the used combination rules, see [10] for their description.

The *Dempster rule* is the normalized conjunctive combination rule of the equation (16) given for two basic belief assignments m_1 and m_2 and for all $X \in G^\Theta$, $X \neq \emptyset$ by:

$$m_{\mathrm{DS}}(X) = \frac{1}{1-k} \sum_{A \cap B = X} m_1(A) m_2(B). \tag{35}$$

where k is either $m_{\mathrm{c}}(\emptyset)$ or the sum of the masses of the elements of $\emptyset$ equivalence class for D^Θ.

The *Yager rule* transfers the global conflict on the total ignorance Θ:

$$m_{\mathrm{Y}}(X) = \begin{cases} m_{\mathrm{c}}(X) & \text{if } X \in 2^\Theta \setminus \{\emptyset, \Theta\} \\ m_{\mathrm{c}}(\Theta) + m_{\mathrm{c}}(\emptyset) & \text{if } X = \Theta \\ 0 & \text{if } X = \emptyset \end{cases} \tag{36}$$

The disjunctive combination rule is given for two basic belief assignments m_1 and m_2 and for all $X \in G^\Theta$ by:

$$m_{\mathrm{Dis}}(X) = \sum_{A \cup B = X} m_1(A) m_2(B). \tag{37}$$

The Dubois and Prade rule is given for two basic belief assignments m_1 and m_2 and for all $X \in G^\Theta$, $X \neq \emptyset$ by:

$$m_{\mathrm{DP}}(X) = \sum_{A \cap B = X} m_1(A) m_2(B) + \sum_{\substack{A \cup B = X \\ A \cap B \equiv \emptyset}} m_1(A) m_2(B). \tag{38}$$

The PCR rule is given for two basic belief assignments m_1 and m_2 and for all $X \in G^\Theta$, $X \neq \emptyset$ by:

$$m_{\mathrm{PCR}}(X) = m_{\mathrm{c}}(X) + \sum_{\substack{Y \in G^\Theta, \\ X \cap Y \equiv \emptyset}} \left(\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)} \right), \tag{39}$$

The principle is very simple: compute the degree of specificity of the bba's you want combine, then calculate the degree of specificity obtained on the bba after the chosen combination rule. The degree of specificity can be compared to the degrees of specificity of the combined bba's.

In the following example given in the table III we combine two Bayesian bba's with the discernment frame $\Theta = \{\theta_1, \theta_2, \theta_3\}$. Both bba's are very contradictory. The values are rounded up. The first approach gives the same degree of specificity than the second one except for the rules m_{Dis}, m_{DP} and m_{Y}. We observe that the smallest degree of specificity is obtained for the conjunctive rule because of the accumulated mass on the empty set not considered in the calculus of the degree. The highest degree of specificity is reached for the Yager rule, for the same reason. That is the only rule given a degree of specificity superior to $\delta_S(m_1)$ and to $\delta_S(m_2)$. The second approach shows well the loss of specificity with the rules m_{Dis}, m_{Y} and m_{DP} having a cautious comportment. With the example, the degree of specificity obtained by the combination rules are almost all inferior to $\delta_S(m_1)$ and to $\delta_S(m_2)$, because the bba's are very conflicting. If the degrees of specificity of the rule such as m_{DS} and m_{PCR} are superior

Table III
DEGREES OF SPECIFICITY FOR COMBINATION RULES ON BAYESIAN BBA'S.

	m_1	m_2	m_{c}	m_{DS}	m_{Y}	m_{Dis}	m_{DP}	m_{PCR}
$\emptyset$	0	0	0.76	0	0	0	0	0
θ_1	0.6	0.2	0.12	0.50	0.12	0.12	0.12	0.43
θ_2	0.1	0.6	0.06	0.25	0.06	0.06	0.06	0.37
θ_3	0.3	0.2	0.06	0.25	0.06	0.06	0.06	0.20
$\theta_1 \cup \theta_2$	0	0	0	0	0	0.38	0.38	0
$\theta_1 \cup \theta_3$	0	0	0	0	0	0.18	0.18	0
$\theta_2 \cup \theta_3$	0	0	0	0	0	0.20	0.20	0
Θ	0	0	0	0	0.76	0	0	0
m_s 1-	m_{θ_1}	m_{θ_2}	m_{θ_1}	m_{θ_1}	m_Θ	$m_{\theta_1 \cup \theta_2}$	$m_{\theta_1 \cup \theta_2}$	m_{θ_1}
m_s 2-	m_{θ_1}	m_{θ_2}	m_{θ_1}	m_{θ_1}	m_{θ_1}	m_{θ_1}	m_{θ_1}	m_{θ_1}
δ_S 1-	0.639	0.655	0.176	0.567	0.857	0.619	0.619	0.497
δ_S 2-	0.639	0.655	0.176	0.567	0.457	0.478	0.478	0.497

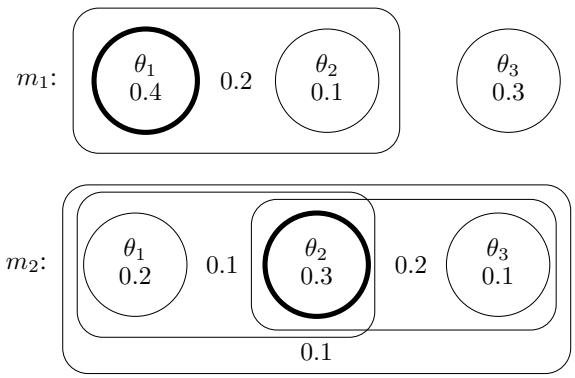
Table IV
DEGREES OF SPECIFICITY FOR COMBINATION RULES ON NON-BAYESIAN BBA'S.

	m_1	m_2	m_{c}	m_{DS}	m_{Y}	m_{Dis}	m_{DP}	m_{PCR}
$\emptyset$	0	0	0.47	0	0	0	0	0
θ_1	0.4	0.2	0.2	0.377	0.2	0.08	0.2	0.39
θ_2	0.1	0.3	0.17	0.321	0.17	0.03	0.17	0.28
θ_3	0.3	0.1	0.12	0.226	0.12	0.03	0.12	0.24
$\theta_1 \cup \theta_2$	0.2	0.1	0.04	0.076	0.04	0.31	0.18	0.06
$\theta_1 \cup \theta_3$	0	0	0	0	0	0.1	0.1	0
$\theta_2 \cup \theta_3$	0	0.2	0	0	0	0.18	0.1	0.03
Θ	0	0.1	0	0	0.47	0.27	0.13	0
m_s 1-	m_{θ_1}	m_{θ_2}	m_{θ_1}	m_{θ_1}	m_{θ_1}	$m_{\theta_1 \cup \theta_2}$	m_{θ_1}	m_{θ_1}
m_s 2-	m_{θ_1}	m_{θ_2}	m_{θ_1}	m_{θ_1}	m_{θ_1}	m_{θ_1}	m_{θ_1}	m_{θ_1}
δ_S 1-	0.553	0.522	0.336	0.488	0.389	0.609	0.428	0.497
δ_S 2-	0.553	0.522	0.336	0.488	0.389	0.456	0.428	0.497

to $\delta_S(m_1)$ and to $\delta_S(m_2)$, that means that the bba's are not in conflict.

Let's consider now a simple non-Bayesian example in table IV.

Figure 4. Two non-Bayesian bba's



VI. CONCLUSION

First, we propose in this article a reflection on the measures on uncertainty in the theory of belief functions. A lot of measures have been proposed to quantify different kind of uncertainty such as the specificity - very linked to the imprecision - and the discord. The discord, we do not have to confuse with the conflict, is for us a contradiction of a source (giving information with a bba in the theory of belief functions) with oneself. We distinguish the contradiction and the conflict that is the contradiction between 2 or more bba's. We introduce a measure of contradiction for a bba based on the weighted average of the conflict between the bba and the categorical bba's of the focal elements.

The previous proposed specificity or non-specificity measures are not defined on the same space. Therefore that is difficult to compare them. That is the reason why we propose the use of degree of uncertainty. Moreover these measures give some counter-intuitive results on Bayesian bba's. We propose a degree of specificity based on the distance between a mass and its most specific associated mass that we introduce. This most specific associated mass can be obtained by two ways and give the nearest categorical bba for a given bba. We propose also to use the degree of specificity in order to measure the specificity of a fusion rule. That is a tool to compare and evaluate the several combination rules given in the theory of belief functions.

Acknowledgments

The authors want to thanks BREST METROPOLE OCÉANE and ENSTA BRETAGNE for funding this collaboration and providing them an excellent research environment during spring 2010.

REFERENCES

[1] A.P. Dempster, "Upper and Lower probabilities induced by a multivalued mapping", *Annals of Mathematical Statistics*, vol. 83, pp. 325-339, 1967.
[2] D. Dubois and H. Prade, "A note on measures of specificity for fuzzy sets", *International Journal of Genera System*, vol. 10, pp. 279-283, 1985.
[3] D. Dubois and H. Prade, "A set-theoretic view on belief functions : logical operations and approximations by fuzzy sets", *International Journal of General Systems*, vol. 12, pp. 193-226, 1986.
[4] M.C. Florea, and E. Bossé, "Crisis Management using Dempster Shafer Theory: Using dissimilarity measures to characterize sources' reliability", in *C3I for Crisis, Emergency and Consequence Management*, Bucharest, Romania, 2009.
[5] R.V.L. Hartley, "Transmission of information", *The Bell System Technical Journal*, vol. 7(3), pp. 535-563, 1928.
[6] A.-L. Jousselme, D. Grenier and E. Bossé, "A new distance between two bodies of evidence", *Information Fusion*, vol. 2, pp. 91-101, 2001.
[7] T. George and N.R. Pal, "Quantification of conflict in Dempster-Shafer framework: a new approach", *International Journal of General Systems*, vol. 24(4), pp. 407-423, 1996,
[8] G.J. Klir, "Measures of uncertainty in the Dempster-Shafer theory of evidence", *Advances in the Dempster-Shafer Theory of Evidence*, John Wiley and Sons, New York, R.R. Yager and M. Fedrizzi and J. Kacprzyk edition, pp. 35-49, 1994.

[9] W. Liu, "Analyzing the degree of conflict among belief functions", *Artificial Intelligence*, vol. 170, pp. 909-924, 2006.
[10] A. Martin and C. Osswald, "Toward a combination rule to deal with partial conflict and specificity in belief functions theory", *International Conference on Information Fusion*, Québec, Canada, 9-12 July 2007.
[11] A. Martin, A.-L. Jousselme and C. Osswald, "Conflict measure for the discounting operation on belief functions," *International Conference on Information Fusion*, Cologne, Germany, 30 June-3 July 2008.
[12] C. Osswald and A. Martin, "Understanding the large family of Dempster-Shafer theory's fusion operators - a decision-based measure", *International Conference on Information Fusion*, Florence, Italy, 10-13 July 2006.
[13] A. Ramer, "Uniqueness of information measure in the theory of evidence", *Fuzzy Sets and Systems*, vol. 24, pp. 183-196, 1987.
[14] B. Ristic and Ph. Smets, "The TBM global distance measure for association of uncertain combat ID declarations", *Information fusion*, vol. 7(3), pp. 276-284, 2006.
[15] G. Shafer, *A mathematical theory of evidence*. Location: Princeton University Press, 1976.
[16] C.E. Shannon, "A mathematical theory of communication", *Bell System Technical Journal*, vol. 27, pp. 379-423, 1948.
[17] F. Smarandache and J. Dezert, *Applications and Advances of DSmT for Information Fusion*. American Research Press Rehoboth, 2004.
[18] Ph. Smets, "Constructing the pignistic probability function in a context of uncertainty", in *Uncertainty in Artificial Intelligence*, vol. 5, pp. 29–39, 1990.
[19] Ph. Smets, "Analyzing the combination of conflicting belief functions", *Information Fusion*, vol. 8, no. 4, pp. 387-412, 2007.
[20] M.J. Wierman, "Measuring Conflict in Evidence Theory", *IFSA World Congress and 20th NAFIPS International Conference*, vol. 3(21), pp. 1741-1745, 2001.
[21] R.R. Yager, "Entropy and Specificity in a Mathematical Theory of Evidence", *International Journal of General Systems*, vol. 9, pp. 249-260, 1983.

Evidential Reasoning for Multi-Criteria Analysis Based on DSmT-AHP

Jean Dezert
Jean-Marc Tacnet

Originally published as Dezert J., Tacnet J.-M., *Evidential Reasoning for Multi-Criteria Analysis based on DSmT-AHP,* ISAHP 2011, Italy, June 2011, and reprinted with permission.

Abstract

In this paper, we present an extension of the multi-criteria decision making based on the Analytic Hierarchy Process (AHP) which incorporates uncertain knowledge matrices for generating basic belief assignments (bba's). The combination of priority vectors corresponding to bba's related to each (sub)-criterion is performed using the Proportional Conflict Redistribution rule no. 5 proposed in Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning. The method presented here, called DSmT-AHP, is illustrated on very simple examples.

Keywords: Analytic Hierarchy Process (AHP), Dezert-Smarandache Theory (DSmT), Evidential Reasoning, Information Fusion, Decision Making, Multi-Criteria Analysis (MCA).

I. Introduction

In many real-life contexts, decisions are based on imperfect information provided by several more or less reliable and conflicting sources. Several recent developments have tried to introduced belief function theory [24] into the AHP framework. A first attempt has been done to consider imprecise evaluations of alternatives using the Dempster-Shafer theory and the Dempster rule [2]. However, the classical fusion rules such as Dempster rule is known to poorly take conflict into account. Another new framework called *ER-MCDA* has been proposed to mix a multi-criteria decision making method called Analytic Hierarchy Process (AHP) and uncertainty theories including fuzzy sets, possibility and belief function theories [28]. The principle of the $ER-MCDA$ methodology is to use AHP to analyze the decision problem and to replace the aggregation step by two successive fusion processes [29]. Its main objective is to take into account both information imperfection, source reliability and conflict. When doing this, an important problem occurs since the importance of criteria is a different concept than the classical reliability concept developped and used in the belief theory context. This article presents a new development related to the use of fusion in the context of multicriteria decision analysis, focusing on the special case of AHP. First, we present a method called DSmT-AHP which is an adapted version of AHP allowing to consider imprecise and uncertain evaluation of alternatives. Secondly, we describe a new fusion rule that has been developped on the basis of the PCR rule proposed in the context of DSmT Theory to implement this method. This new rule is also used in the context of the ER-MCDA method.

II. DSmT-AHP approach

We present briefly the extension of Saaty's AHP method [17] using an aggregation method developed in Dezert-Smarandache Theory (DSmT) framework [25] of evidential reasoning (ER), able to make a difference between importance of criteria, uncertainty related to the evaluations of criteria and reliability of the different sources. This method has been introduced for the first time in [8] and applied for risk prevention of natural hazards in mountains in [29]. Before explaining the principle of DSmT-AHP, it is necessary to recall some basics of DSmT at first to make the paper self-consistent for readers not familiar with this theoretical framework. All basis with many detailed examples can be obtained freely by downloading the three volumes given in [25] from the web.

A. DSmT basics

Let start with $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$ be a finite set of n elements assumed to be exhaustive. Θ corresponds to the frame of discernment of the problem under consideration. In general, we assume that elements of Θ are non exclusive in order to deal with vague/fuzzy and relative concepts [25], Vol. 2. This is the so-called free-DSm model. In DSmT, there is no need to work on a refined frame consisting in a discrete finite set of exclusive and exhaustive hypotheses (referred as Shafer's model), because DSm rules of combination work for any models of the frame. The hyper-power set D^Θ is defined as the set of all propositions built from elements of Θ with $\cup$ and $\cap$, see [25], Vol. 1 for examples. A (quantitative) basic belief assignment (bba) expressing the belief committed to the elements of D^Θ by a given source is a mapping $m(\cdot)$: $D^\Theta \rightarrow [0,1]$ such that: $m(\emptyset) = 0$ and $\sum_{A \in D^\Theta} m(A) = 1$. Elements $A \in D^\Theta$ having $m(A) > 0$ are called *focal elements* of $m(.)$. The credibility

and plausibility functions are defined in almost the same manner as in Dempster-Shafer Theory (DST) [24] except that 2^Θ is replaced by D^Θ in their definitions. In DSmT, the Proportional Conflict Redistribution Rule no. 5 (PCR5) is used generally to combine bba's. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. For simplicity, we work in the power set 2^Θ since most of readers are usually already familiar with this fusion space. Let's $m_1(.)$ and $m_2(.)$ be two independentbba's, then the PCR5 rule is defined as follows: $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)} \right] \tag{1}$$

where all denominators in (1) are different from zero. If a denominator is zero, that fraction is discarded. All propositions/sets are in a canonical form. A variant of (1), called PCR6, for combining $s > 2$ sources and for working in other fusion spaces (hyper-power sets or super power-sets) is presented in [25]. Additional properties of PCR5 can be found in [7]. Extension of PCR5 for combining qualitative bba's can be found in [25], Vol. 2 & 3.

As a simple example, let's consider two bba's $m_1(.)$ and $m_2(.)$, $A \cap B = \emptyset$ for the model of Θ, and $m_1(A) = 0.6$ and $m_2(B) = 0.3$. With PCR5 the partial conflicting mass $m_1(A)m_2(B) = 0.6 \cdot 0.3 = 0.18$ is redistributed to A and B only with respect to the following proportions respectively: $x_A = 0.12$ and $x_B = 0.06$ because

$$\frac{x_A}{m_1(A)} = \frac{x_B}{m_2(B)} = \frac{m_1(A)m_2(B)}{m_1(A) + m_2(B)} = \frac{0.18}{0.9} = 0.2$$

B. DSmT-AHP approach

DSmT-AHP aimed to perform a similar purpose as AHP [16], [17], SMART [30] or DS/AHP [2], [4], etc. that is to find the preferences rankings of the decision alternatives (DA), or groups of DA. DSmT-AHP approach consists in three steps:

- Step 1: We extend the construction of the matrix for taking into account the partial uncertainty (disjunctions) between possible alternatives. If no comparison is available between elements, then the corresponding elements in the matrix is zero. Each bba related to each (sub-) criterion is the normalized eigenvector associated with the largest eigenvalue of the "uncertain" knowledge matrix (as done in standard AHP approach).
- Step 2: We use the DSmT fusion rules, typically the PCR5 rule, to combine bba's drawn from step 1 to get a final Multi-Criteria Decision-Making (MCDM) priority ranking. This fusion step must take into account the different importances (if any) of criteria as it will be explained in the sequel.
- Step 3: Decision-making can be done based either on the maximum of belief (also called credibility), or on the maximum of the plausibility of Decision alternatives (DA), as well as on the maximum of the approximate subjective probability of DA obtained by different probabilistic transformations like the Pignistic, DSmP, or Sudano's transformations [25], Vol. 2.

The MCDM problem deals with several criteria having different importances and the classical fusion rules cannot be applied directly as in step 2. In AHP, the fusion is done from the product of the bba's matrix with the weighting vector of criteria. Such AHP fusion is nothing but a simple componentwise weighted average of bba's and it doesn't actually process efficiently the conflicting information between the sources. It doesn't preserve the neutrality of a full ignorant source in the fusion. To palliate these problems, we recall the new solution for combining sources of different importances proposed in [26]. The reliability is an objective property of a source, whereas the importance of a source is a subjective characteristic expressed by the fusion system designer. The reliability of a source represents its ability to provide the correct assessment/solution of the given problem. It is characterized by a discounting reliability factor, usually denoted α in $[0, 1]$, which should be estimated from statistics when available, or by other techniques [12]. The reliability can be context-dependent. By convention, we usually take $\alpha = 1$ when the source is fully reliable and $\alpha = 0$ if the source is totally unreliable. The reliability of a source is usually taken into account with Shafer's discounting method [24] defined by:

$$\begin{cases} m_\alpha(X) = \alpha \cdot m(X), & \text{for } X \neq \Theta \\ m_\alpha(\Theta) = \alpha \cdot m(\Theta) + (1 - \alpha) \end{cases} \tag{2}$$

The importance of a source is not the same as its reliability and we characterize it by an importance factor β in $[0, 1]$ which represents somehow the weight of importance granted to the source by the fusion system designer. The choice of β is usually not related with the reliability of the source and can be chosen to any value in $[0, 1]$ by the designer for his/her own reason. By convention, the fusion system designer will take $\beta = 1$ when he/she wants to grant the maximal importance of the source in the fusion process, and will take $\beta = 0$ if no importance at all is granted to this source in the fusion process. The fusion designer must be able to deal with importance factors in a different way than with reliability factors since they correspond to distinct properties associated with a source of information. The importance of a source is particularly crucial in hierarchical multi-criteria decision making problems, specially in the AHP [17], [21]. That's why it is primordial to show how the importance

can be efficiently managed in evidential reasoning approaches. To take into account the importance of the sources we use the dual of Shafer's discounting approach for reliabilities of sources as originally introduced briefly by Tacnet in [25], Vol.3, Chap. 23, p. 613. It consists to define the importance discounting with respect to the empty set rather than the total ignorance Θ (as done with Shafer's discounting). Such new discounting deals easily with sources of different importances and is very simple to use. Mathematically, we define the importance discounting of a source $m(.)$ having the importance factor β in $[0,1]$ by:

$$\begin{cases} m_\beta(X) = \beta \cdot m(X), \quad \text{for } X \neq \emptyset \\ m_\beta(\emptyset) = \beta \cdot m(\emptyset) + (1-\beta) \end{cases} \tag{3}$$

Here we allow to deal with non-normal bba since $m_\beta(\emptyset) \geq 0$ as suggested by Smets in [27]. This new discounting preserves the specificity of the primary information since all focal elements are discounted with same importance factor. Here we use the positive mass of the empty set as an intermediate/preliminary step of the fusion process. Clearly when $\beta = 1$ is chosen by the fusion designer, it will mean that the source must take its full importance in the fusion process and so the original bba $m(.)$ is kept unchanged. If the fusion designer takes $\beta = 0$, one will deal with $m_\beta(\emptyset) = 1$ which is interpreted as a fully non important source. $m(\emptyset) > 0$ is not interpreted as the mass committed to some conflicting information (classical interpretation), nor as the mass committed to unknown elements when working with the open-world assumption (Smets interpretation), but only as the mass of the discounted importance of a source in this particular context. Based on this discounting, one adapts PCR5 (or PCR6) rule for $N \geq 2$ discounted bba's $m_{\beta,i}(.)$, $i = 1, 2, \ldots N$ by considering the following extension, denoted PCR5$_\emptyset$, defined by: $\forall X \in 2^\Theta$

$$m_{PCR5_\emptyset}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)} \right] \tag{4}$$

A similar extension can be done for PCR5 and PCR6 formulas for $N > 2$ sources given in [25], Vol. 2. A detailed presentation of this technique with several examples has been done in [26] and thus it is not reported here due to space limitation constraints. The difference between eqs. (1) and (4) is that $m_{PCR5}(\emptyset) = 0$ whereas $m_{PCR5_\emptyset}(\emptyset) \geq 0$. Since we usually work with normal bba's for decision making support, the combined bba will be normalized. In the AHP context, the importance factors correspond to the components of the normalized eigenvector $\mathbf{w}$. It is important to note that such importance discounting method cannot be used in DST when using Dempster-Shafer's rule of combination because this rule is not responding to the discounting of sources towards the empty set (see Theorem 1 in [26] for proof).

We have shown how the reliability and importance of sources can be taken into account in the fusion process separately. The possibility to take them into account jointly is more difficult, because in general the reliability and importance discounting approaches do not commute, but when $\alpha_i = \beta_i = 1$. Indeed, it can be easily verified that $m_{\alpha_i,\beta_i}(.) \neq m_{\beta_i,\alpha_i}(.)$ whenever $\alpha_i \neq 1$ and $\beta_i \neq 1$. $m_{\alpha_i,\beta_i}(.)$ denotes the reliability discounting of $m_i(.)$ by α_i followed by the importance discounting of $m_{\alpha_i}(.)$ by β_i which explains the notation α_i, β_i used in index. Similarly, $m_{\beta_i,\alpha_i}(.)$ denotes the importance discounting of $m_i(.)$ by β_i followed by the reliability discounting of $m_{\beta_i}(.)$ by α_i. In order to deal both with reliabilities and importances factors and because of the non commutativity of these discountings, we have proposed two methods in [26] to proceed the fusion of the sources in a three-steps process as follows:

• Method 1: Step 1: Apply reliability and then importance discountings to get $m_{\alpha_i,\beta_i}(.)$, $i = 1, \ldots, s$ and combine them with $PCR5_\emptyset$ or $PCR6_\emptyset$ and normalize the resulting bba; Step 2: Apply importance and then reliability discountings to get $m_{\beta_i,\alpha_i}(.)$, $i = 1, \ldots, s$ and combine them with $PCR5_\emptyset$ or $PCR6_\emptyset$ and normalize the resulting bba; Step 3 (mixing/averaging): Combine the resulting bba's of Steps 1 and 2 using the arithmetic mean operator[1].

• Method 2: Another simplest method which could be useful for intermediate traceability in some applications would consist to change Steps 1 & 2 by Step 1': Apply reliability discounting only to get $m_{\alpha_i}(.)$ and combine them with PCR5 or PCR6; Step 2': Apply importance discounting only to get $m_{\beta_i}(.)$ and combine them with $PCR5_\emptyset$ or $PCR6_\emptyset$ and normalize the result; Step 3' same as Step 3. Due to space limitation, only Method 1 is briefly illustrated in the following simple example.

[1]Other combination rules could be used also like PCR5 or PCR6, etc., but we don't see solid justification to use them again and they require more computations than the simple arithmetic mean.

C. A simple example

Let's consider a set of three cars $\Theta = \{A, B, C\}$ with Shafer's model, and only two criteria C1=Fuel Economy, C2=Reliability. Let's assume that with respect to each criterion the following "uncertain" preferences matrices are given:

$$\mathbf{M}(C1) = \left[\begin{array}{c|ccc} & A & B\cup C & \Theta \\ \hline A & 1 & 0 & 1/3 \\ B\cup C & 0 & 1 & 2 \\ \Theta & 3 & 1/2 & 1 \end{array}\right], \qquad \mathbf{M}(C2) = \left[\begin{array}{c|cccc} & A & B & A\cup C & B\cup C \\ \hline A & 1 & 2 & 4 & 3 \\ B & 1/2 & 1 & 1/2 & 1/5 \\ A\cup C & 1/4 & 2 & 1 & 0 \\ B\cup C & 1/3 & 5 & 0 & 1 \end{array}\right]$$

• **Step 1 (bba's generation)**: Applying AHP method, one gets the following priority vectors $\mathbf{w}(C1) \approx [0.0889\ 0.5337\ 0.3774]'$ and $\mathbf{w}(C2) \approx [0.5002\ 0.1208\ 0.1222\ 0.2568]'$ which are identified with the bba's $m_{C1}(.)$ and $m_{C2}(.)$ as follows: $m_{C1}(A) = 0.0889$, $m_{C1}(B \cup C) = 0.5337$, $m_{C1}(A \cup B \cup C) = 0.3774$ and $m_{C2}(A) = 0.5002$, $m_{C2}(B) = 0.1208$, $m_{C2}(A \cup C) = 0.1222$ and $m_{C2}(B \cup C) = 0.2568$.

• **Step 2 (Fusion)**: When the two criteria have the same full importance in the hierarchy they are fused with one of the classical fusion rules. In [4] Beynon proposed to use Dempster's rule. Here we propose to use the PCR5 fusion rule since it is known to have a better ability to deal efficiently with possibly highly conflicting sources of evidences [25], Vol. 2. With PCR5, one gets:

Elem. of 2^Θ	$m_{C1}(.)$	$m_{C2}(.)$	$m_{PCR5}(.)$
$\emptyset$	0	0	0
A	0.0889	0.5002	0.3837
B	0	0	0.1162
$A \cup B$	0	0.1208	0
C	0	0	0.0652
$A \cup C$	0	0.1222	0.0461
$B \cup C$	0.5337	0.2568	0.3887
$A \cup B \cup C$	0.3774	0	0

• **Step 3 (Decision-making)**: A final decision based on $m_{PCR5}(.)$ must be taken. Usually, the decision-maker (DM) is concerned with a single choice among the elements of Θ. Many decision-making approaches are possible depending on the risk the DM is ready to take. A pessimistic DM will choose the singleton of Θ giving the maximum of credibility whereas an optimistic DM will choose the element having the maximum of plausibility. A fair attitude consists usually in choosing the maximum of approximate subjective probability of elements of Θ. The result however is very dependent on the probabilistic transformation (Pignistic, DSmP, Sudano's, etc) [25], Vol. 2. We recall[2] that the credibility Bel(.), the pignistic probability BetP(.) and the plausibility Pl(.) functions satisfy for any $A \in 2^\Theta$ the following inequality:

$$\text{Bel}(A) \leq \text{BetP}(A) \leq \text{Pl}(A) \tag{5}$$

where $\text{Bel}(A)$, $\text{Pl}(A)$ and $\text{BetP}(A)$ are defined from any bba $m(.)$ by the following formulas:

$$\text{Bel}(A) = \sum_{\substack{B\subseteq A \\ B\in 2^\Theta}} m(B) \qquad \text{and} \qquad \text{Pl}(A) = \sum_{\substack{B\cap A\neq\emptyset \\ B\in 2^\Theta}} m(B) \tag{6}$$

$$\text{BetP}(A) = \sum_{Y\in 2^\Theta} \frac{|Y\cap A|}{|Y|} m(Y) \tag{7}$$

and where $|Y|$ denotes the cardinality of Y. By convention one takes $|\emptyset|/|\emptyset| = 1$.
Below are the values of the credibility, the pignistic probability and the plausibility of A, B and C:

Elem. of Θ	Bel(.)	BetP(.)	Pl(.)
A	0.3837	0.4068	0.4298
B	0.1162	0.3105	0.5049
C	0.0652	0.2826	0.5000

The car A will be preferred with the pessimistic (based on max of Bel(.)) or pignistic attitudes, whereas the car B will be preferred if an optimistic attitude is adopted since one has $\text{Pl}(B) > \text{Pl}(C) > \text{Pl}(A)$.

[2]In this example, we work with Shafer's model for the frame Θ so that $D^\Theta = 2^\Theta$.

Now, if we assume that C2 (the reliability) is three times more important than C1 (fuel economy) so that the preference matrix between these two criteria is given by:

$$\mathbf{M} = \begin{bmatrix} 1/1 & 1/3 \\ 3/1 & 1/1 \end{bmatrix} \approx \begin{bmatrix} 1.0000 & 0.3333 \\ 3.0000 & 1.0000 \end{bmatrix}$$

Its normalized principal eigenvector is $\mathbf{w} = [0.2500\ 0.7500]'$ which indicates that C2 is indeed three times more important than C1 as expressed in the prior DM preferences for ranking criteria. $\mathbf{w} = [w_1\ w_2]'$ can also be obtained directly by solving the algebraic system of equations $w_2 = 3w_1$ and $w_1 + w_2 = 1$ with $w_1, w_2 \in [0, 1]$. If we apply the importance discounting with $\beta_1 = w_1 = 0.25$ and $\beta_2 = w_2 = 0.75$, one gets the following discounted bba's

Elem. of 2^Θ	$m_{\beta_1,C1}(.)$	$m_{\beta_2,C2}(.)$
$\emptyset$	0.7500	0.2500
A	0.0222	0.3751
B	0	0
$A \cup B$	0	0.0906
C	0	0
$A \cup C$	0	0.0917
$B \cup C$	0.1334	0.1926
$A \cup B \cup C$	0.0944	0

With the PCR5$_\emptyset$ fusion of the sources $m_{\beta_1,C1}(.)$ and $m_{\beta_2,C2}(.)$, one gets the results in the table. For decision-making support, one prefers to work with normal bba's. Therefore $m_{PCR5_\emptyset}(.)$ is normalized by redistributing back $m_{PCR5_\emptyset}(\emptyset)$ proportionally to the masses of other focal elements as shown in the right column of the next table.

Elem. of 2^Θ	$m_{PCR5_\emptyset}(.)$	$m^{\text{normalized}}_{PCR5_\emptyset}(.)$
$\emptyset$	0.6558	0
A	0.1794	0.5213
B	0.0121	0.0351
$A \cup B$	0.0159	0.0461
C	0.0122	0.0355
$A \cup C$	0.0161	0.0469
$B \cup C$	0.1020	0.2963
$A \cup B \cup C$	0.0065	0.0188

If all sources have the same full importances (i.e. all β_i=1), then $m_{PCR5_\emptyset}(.) = m_{PCR5}(.)$ which is normal because in such case $m_{\beta_i=1,Ci}(.) = m_{Ci}(.)$. From $m^{\text{normalized}}_{PCR5_\emptyset}(.)$ one can easily compute the credibility Bel(.), the pignistic probability BetP(.) or the plausibility Pl(.) for each element of Θ for decision-making. In this example one gets:

Elem. of Θ	Bel(.)	BetP(.)	Pl(.)
A	0.5213	0.5741	0.6331
B	0.0351	0.2126	0.3963
C	0.0355	0.2134	0.3974

If the classical AHP *fusion* method (i.e. weighted arithmetic mean) is used directly with bba's $m_{C1}(.)$ and $m_{C2}(.)$, one gets:

$$m_{AHP}(.) = \begin{bmatrix} 0 & 0 \\ 0.0889 & 0.5002 \\ 0 & 0 \\ 0 & 0.1208 \\ 0 & 0 \\ 0 & 0.1222 \\ 0.5337 & 0.2568 \\ 0.3774 & 0 \end{bmatrix} \times \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.3974 \\ 0 \\ 0.0906 \\ 0 \\ 0.0917 \\ 0.3260 \\ 0.0944 \end{bmatrix}$$

which would have provided the following result for decision-making

Elem. of Θ	Bel(.)	BetP(.)	Pl(.)
A	0.3974	0.5200	0.6741
B	0	0.2398	0.5110
C	0	0.2403	0.5121

In this very simple example, one sees that the importance discounting technique coupled with PCR5-based fusion rule (what we call the DSmT-AHP approach) will suggest, as with classical AHP, to choose the alternative A since the car A has a bigger credibility (as well as a bigger pignistic probability and plausibility) than cars B or C. It is however worth to note that the values of Bel(.), BetP(.) and Pl(.) obtained by both methods are slightly different. The difference in results can have a strong impact in practice in the final result for example if the costs of vehicles have also to be included in the final decision. Note

also that the uncertainties $U(X) = \text{Pl}(X) - \text{Bel}(X)$ of all alternatives $X = A, B, C$ have been seriously diminished when using DSmT-AHP with respect to what we obtain with classical AHP as seen in the following table. The uncertainty reduction is a nice expected property specially important for decision-making support.

Elem. of Θ	$U(.)$ with AHP	$U(.)$ with DSmT-AHP
A	0.2767	0.1118
B	0.5110	0.3612
C	0.5121	0.3619

III. Conclusions

In this paper, we have presented a new method for Multi-Criteria Decision-Making (MCDM) and Multi-Criteria Group Decision-Making (MCGDM) based on the combination of AHP method developed by Saaty and DSmT. The AHP method allows to build bba's from DM preferences of solutions which are established with respect to several criteria. The DSmT allows to aggregate efficiently the (possibly highly conflicting) bba's based on each criterion. This DSmT-AHP method allows to take into account also the different importances of the criteria and/or of the different members of the decision-makers group. The application of this DSmT-AHP approach and specially the new fusion rule has been introduced in an application case for the prevention of natural hazards in mountains and snow avalanches, see [25], Vol.3, Chap. 23, and [29].

References

[1] J. Barzilai, F.A. Lootsma, *Power relations and group aggregation in the multiplicative AHP and SMART*, J. of MCDA, Vol. 6, pp. 155–165, 1997.

[2] M. Beynon, B. Curry, P.H. Morgan, *The Dempster-Shafer theory of evidence: An alternative approach to multicriteria decision modelling*, Omega, Vol. 28, No. 1, pp. 37–50, 2000.

[3] M. Beynon, D. Cosker, D. Marshall, *An expert system for multi-criteria decision making using Dempster-Shafer theory*, Expert Syst. with Appli., Vol. 20, No. 4, pp. 357–367, 2001.

[4] M. Beynon, *DS/AHP method: A mathematical analysis, including an understanding of uncertainty*, Eur. J. of Oper. Res., Vol. 140, pp. 148–164, 2002.

[5] M. Beynon, *Understanding local ignorance and non-specificity within the DS/AHP method of multi-criteria decision making*, Eur. J. of Oper. Res., Vol. 163, pp. 403–417, 2005.

[6] M. Beynon, *A method of aggregation in DS/AHP for group decision-making with non-equivalent importance of individuals in the group*, Comp. & Oper. Res., No. 32, pp. 1881-1896, 2005.

[7] J. Dezert, F. Smarandache, *Non Bayesian conditioning and deconditioning*, International Workshop on Belief Functions, Brest, France, April 2010.

[8] J. Dezert, J.-M. Tacnet, M. Batton-Hubert, F. Smarandache, *Multi-criteria decision making based on DSmT-AHP*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[9] E.H. Forman, S.I. Gass, *The analytical hierarchy process: an exposition*, Oper. Res., Vol. 49, No. 4 pp. 46–487, 2001.

[10] R.D. Holder, *Some Comment on the Analytic Hierarchy Process*, J. of the Oper. Res. Soc., Vol. 41, No. 11, pp. 1073–1076, 1990.

[11] F.A. Lootsma, *Scale sensitivity in the multiplicative AHP and SMART*, J. of MCDA, Vol. 2, pp. 87–110, 1993.

[12] A. Martin, A.-L. Jousselme, C. Osswald, *Conflict measure for the discounting operation on belief functions*, Proc. of Fusion 2008, Cologne, Germany, July 2008.

[13] C. K. Murphy, *Combining belief functions when evidence conflicts*, Dec. Sup. Syst., vol. 29, pp. 1-9, 2000.

[14] J. Perez, *Some comments on Saaty's AHP*, Manag. Sci., Vol. 41, No. 6, pp. 1091–1095, 1995.

[15] J. Perez, J.L. Jimeno, E. Mokotoff, *Another Potential Shortcoming of AHP*, TOP: An Official J. of the Spanish Soc. of Stats and Oper. Res., Vol. 14, No. 1 , June, 2006.

[16] T.L. Saaty, *A scaling method for priorities in hierarchical structures*, J. of Math. Psych., Vol. 15, PP. 59–62, 1977.

[17] T.L. Saaty, *The Analytical Hierarchy Process*, McGraw Hill, New York, 1980.

[18] T.L. Saaty, *An exposition of the AHP in reply to the paper "Remarks on the Analytic Hierarchy Process"*, Manag. Sci., Vol. 36, No. 3, pp. 259–268, 1990.

[19] T.L. Saaty, *Response to Holder's Comments on the Analytic Hierarchy Process*, J. of the Oper. Res. Soc., Vol. 42, No. 10, pp. 909–914, October 1991.

[20] T.L. Saaty, *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World*, RWS Publ., Pittsburgh, PA, USA, 1999.

[21] T.L. Saaty, *Fundamentals of decision making and priority theory with the analytic hierarchy process*, Vol. VI of the AHP series, RWL Publ., Pittsburgh, PA, USA.

[22] T.L. Saaty, *Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors - The Analytic Hierarchy/Network Process*, RACSAM, Vol. 102, No. 2, pp. 251–318, 2008.

[23] S. Schenkerman, *Inducement of nonexistent order by the analytic hierarchy process*, Dec. Sci., Spring 1997.

[24] G. Shafer, *A mathematical theory of evidence*, Princeton Univ. Press, 1976.

[25] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vol. 1-3, ARP, 2004–2009. http://www.gallup.unm.edu/~smarandache/DSmT.htm

[26] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, in Proceedings of Fusion 2010 Int. Conf., Edinburgh, UK, July 26-29th, 2010.

[27] Ph. Smets Ph., *The Combination of Evidence in the Transferable Belief Model*, IEEE Trans. PAMI 12, pp. 447–458, 1990.

[28] J.M. Tacnet, M. Batton-Hubert and J. Dezert, "Information fusion for natural hazards in mountains" in *Advances and applications of DSmT for Information Fusion- Collected works - Volume 3 - Dezert J., Smarandache F. (Eds.) - American Research Press, Rehoboth, USA.*, pp. 565–659, 2009.

[29] J.-M. Tacnet, M. Batton-Hubert, J. Dezert, *A two-step fusion process for multi-criteria decision applied to natural hazards in mountains*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[30] D. Von Winterfeldt, W. Edwards *Decision analysis and behavioral research*, Cambridge Univ. Press, USA, 1986.

Cautious OWA and Evidential Reasoning for Decision Making under Uncertainty

Jean-Marc Tacnet
Jean Dezert

Originally published as Tacnet J.-M., Dezert J., *Cautious OWA and Evidential Reasoning for Decision Making under Uncertainty,* in Proc. Of Fusion 2011 Conf., Chicago, July, 2011, and reprinted with permission.

***Abstract*—To make a decision under certainty, multicriteria decision methods aims to choose, rank or sort alternatives on the basis of quantitative or qualitative criteria and preferences expressed by the decision-makers. However, decision is often done under uncertainty: choosing alternatives can have different consequences depending on the external context (or state of the word). In this paper, a new methodology called Cautious Ordered Weighted Averaging with Evidential Reasoning (COWA-ER) is proposed for decision making under uncertainty to take into account imperfect evaluations of the alternatives and unknown beliefs about groups of the possible states of the world (scenarii). COWA-ER mixes cautiously the principle of Yager's Ordered Weighted Averaging (OWA) approach with the efficient fusion of belief functions proposed in Dezert-Smarandache Theory (DSmT).**

Keywords: fusion, Ordered Weighted Averaging (OWA), DSmT, uncertainty, information imperfection, multi-criteria decision making (MCDM)

I. Introduction

A. Decisions under certainty, risk or uncertainty

Decision making in real-life situations are often difficult multi-criteria problems. In the classical Multi-Criteria Decision Making (MCDM) framework, those decisions consist mainly in choosing, ranking or sorting alternatives, solutions or more generally potential actions [17] on the basis of quantitative or qualitative criteria. Existing methods differs on aggregation principles (total or partial), preferences weighting, and so on. In total aggregation multicriteria decision methods such as Analytic Hierarchy Process (AHP) [19], the result for an alternative is a unique value called synthesis criterion. Possible alternatives (A_i) belonging to a given set $A = \{A_1, A_2, \ldots, A_q\}$ are evaluated according to preferences (represented by weights w_j) expressed by the decision-makers on the different criteria (C_j) (see figure 1).

Decisions are often taken on the basis of imperfect information and knowledge (imprecise, uncertain, incomplete) provided by several more or less reliable sources and depending on the states of the world: decisions can be taken in certain, risky or uncertain environment. In a MCDM context, decision under certainty means that the evaluations of the alternative are independent from the states of the world. In other cases, alternatives may be assessed differently depending on the scenarii that are considered.

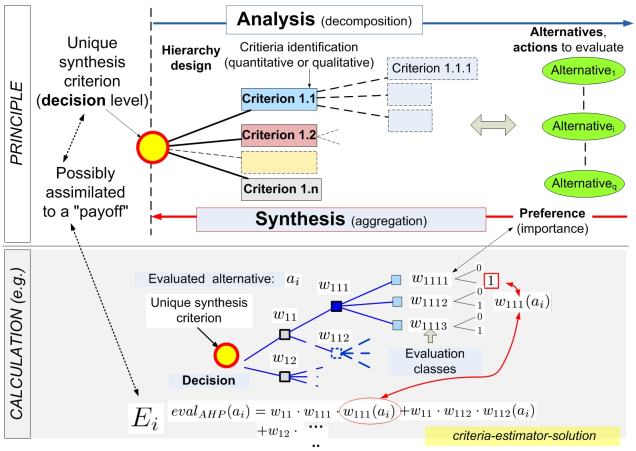


Figure 1. Principle of a multi-criteria decision method based on a total aggregation principle.

In the classical framework of decision theory under uncertainty, Expected Utility Theory (EUT) states that a decision maker chooses between risky or uncertain alternatives or actions by comparing their expected utilities [14]. Let us consider an example of decision under uncertainty (or risk) related to natural hazards management. On the lower parts of torrent catchment basin or an avalanche path, risk analysis consists in evaluating potential damage caused due to the effects of hazard (a phenomenon with an intensity and a frequency) on people and assets at risk. Different strategies (A_i) are possible to protect the exposed areas. For each of them, damage will depend on the different scenarii (S_j) of phenomenon which can be more or less uncertain. An *action* A_i (e.g. building a protection device, a dam) is evaluated through its potential effects r_k to which are associated *utilities* $u(r_k)$ (protection level of people, cost of protection, ...) and *probabilities* $p(r_k)$ (linked to natural events or states of nature S_k). The expected utility $U(a)$ of an action a is estimated through the sum of products of utilities and probabilities of all potential consequences of the action a:

$$U(A_i) = \sum u(r_k) \cdot p(r_k) \qquad (1)$$

When probabilities are known, decision is done under risk. When those probabilities becomes subjective, the prospect theory (subjective expected utility theory - SEUT) [12] can apply :

- *the objective utility* (e.g. cost) $u(r_k)$ is replaced by a subjective function (*value*) denoted $v(u(r_k))$;
- the objective weighting $p(r_k)$ is replaced by a subjective function $\pi(p(r_k))$.

$v(\cdot)$ is the felt subjective value in response of the expected cost of the considered action, and $\pi(\cdot)$ is the felt weighting face to the objective probability of the realisation of the result. Prospect theory shows that the function $v(\cdot)$ is asymmetric: loss causes a negative reaction intensity stronger than the positive reaction caused by the equivalent gain. This corresponds to an aversion to risky choices in the area of earnings and a search of risky choices in the area of loss.

In a MCDM context, information imperfection concerns both the evaluation of the alternatives (in any context of certainty, risk or ignorance) and the uncertainty or lack of knowledge about the possible states of the world. Uncertainty and imprecision in multi-criteria decision models has been early considered [16]. Different kinds of uncertainty can be considered: on the one hand the internal uncertainty is linked to the structure of the model and the judgmental inputs required by the model, on the other hand the external uncertainty refers to lack of knowledge about the consequences about a particular choice.

B. Objectives and goals

Several decision support methods exist to consider both information imperfection, sources heterogeneity, reliability, conflict and the different states of the world when evaluating the alternatives as summarized on figure 2. A more complete review can be found in [28]. Here we just remind some recent examples of methods mixing MCDM approaches and Evidential Reasoning[1](ER).

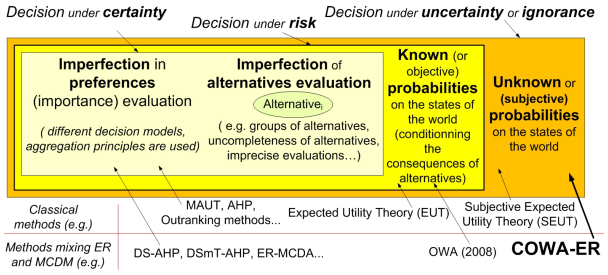

Decision under certainty
Decision under risk
Decision under uncertainty or ignorance
Imperfection in preferences (importance) evaluation
(different decision models, aggregation principles are used)
Imperfection of alternatives evaluation
Alternativei
(e.g. groups of alternatives, uncompleteness of alternatives, imprecise evaluations…)
Known (or objective) probabilities on the states of the world (conditionning the consequences of alternatives)
Unknown or (subjective) probabilities on the states of the world
Classical methods (e.g.)
MAUT, AHP, Outranking methods...
Expected Utility Theory (EUT)
Subjective Expected Utility Theory (SEUT)
Methods mixing ER and MCDM (e.g.)
DS-AHP, DSmT-AHP, ER-MCDA...
OWA (2008)
COWA-ER

Figure 2. Information imperfection in the different decision support methods

- Dempster-Shafer-based AHP (DS-AHP) has introduced a merging of Evidential Reasoning (ER) with Analytic Hierachy Process (AHP) [19] to consider the imprecision and the uncertainty in evaluation of several alternatives. The idea is to consider criteria as sources [1], [3] and derive weights as discounting factors in the fusion process [5];
- Dezert-Smarandache-based (DSmT-AHP) [8] takes into account the partial uncertainty (disjunctions) between possible alternatives and introduces new fusion rules, based on Proportional Conflict Redistribution (PCR) principle, which allow to consider differences between importance and reliability of sources [23];
- ER-MCDA [28], [29] is based on AHP, fuzzy sets theory, possibility theory and belief functions theory too. This method considers both imperfection of criteria evaluations, importance and reliability of sources.

Introducing ignorance and uncertainty in a MCDM process consists in considering that consequences of actions (A_i) depend of the state of nature represented by a finite set $S = \{S_1, S_2, \ldots, S_n\}$. For each state, the MCDM method provides an evaluation C_{ij}. We assume that this evaluation C_{ij} done by the decision maker corresponds to the choice of A_i when S_j occurs with a given (possibly subjective) probability. The evaluation matrix is defined as $C = [C_{ij}]$ where $i = 1, \ldots, q$ and $j = 1, \ldots, n$.

$$\begin{array}{c} \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{array} \begin{array}{c} \begin{array}{cccccc} S_1 & \cdots & S_j & \cdots & S_n \end{array} \\ \begin{pmatrix} C_{11} & \cdots & C_{11} & \cdots & C_{1n} \\ & & \vdots & & \\ C_{i1} & \cdots & C_{ij} & \cdots & C_{in} \\ & & \vdots & & \\ C_{q1} & \cdots & C_{qj} & \cdots & C_{qn} \end{pmatrix} \end{array} = C \qquad (2)$$

Existing methods using evidential reasoning and MCDM have, up to now, focused on the case of imperfect evaluation of alternatives in a context of decision under certainty. In this paper, we propose a new method for decision under uncertainty that mixes MCDM principles, decision under uncertainty principles and evidential reasoning. For this purpose, we propose a framework that considers uncertainty and imperfection for scenarii corresponding to the state of the world.

This paper is organized as follows. In section II, we briefly recall the basis of DSmT. Section III presents two existing methods for MCDM under uncertainty using belief functions theory: DSmT-AHP as an extension of Saaty's multi-criteria decision method AHP, and Yager's Ordered Weighted Averaging (OWA) approach for decision making with belief structures. The contribution of this paper concerns the section IV where we describe an alternative to the classical OWA, called cautious OWA method, where evaluations of alternatives depend on more or less uncertain scenarii. The flexibility and advantages of this COWA method are also discussed. Conclusions and perspectives are given in section V.

[1]Evidential Reasoning refers to the use of belief functions as theoretical background, not to a specific theory of belief functions (BF) aimed for combining, or conditioning BF. Actually, Dempster-Shafer Theory (DST) [21], Dezert-Smarandache Theory (DSmT) [22], and Smets' TBM [25] are different approaches of Evidential Reasoning.

II. Belief functions and DSmT

Dempster-Shafer Theory (DST) [21] offers a powerful mathematical formalism (the belief functions) to model our belief and uncertainty on the possible solutions of a given problem. One of the pillars of DST is Dempster-Shafer rule (DS) of combination of belief functions. The purpose of the development of Dezert-Smarandache Theory (DSmT) [22] is to overcome the limitations of DST by proposing new underlying models for the frames of discernment in order to fit better with the nature of real problems, and new combination and conditioning rules for circumventing problems with DS rule specially when the sources to combine are highly conflicting. In DSmT, the elements θ_i, $i = 1, 2, \ldots, n$ of a given frame Θ are not necessarily exclusive, and there is no restriction on θ_i but their exhaustivity. Some integrity constraints (if any) can be include in the underlying model of the frame. Instead of working in power-set 2^Θ, we classically work on hyper-power set D^Θ (Dedekind's lattice) - see [22], Vol.1 for details and examples. A (generalized) basic belief assignment (bba) given by a source of evidence is a mapping $m : D^\Theta \to [0, 1]$ such that

$$m(\emptyset) = 0 \qquad \text{and} \qquad \sum_{A \in D^\Theta} m(A) = 1 \tag{3}$$

The *generalized* credibility and plausibility functions are defined in almost the same manner as within DST, i.e.

$$\text{Bel}(A) = \sum_{\substack{B \subseteq A \\ B \in D^\Theta}} m(B) \quad \text{and} \quad \text{Pl}(A) = \sum_{\substack{B \cap A \neq \emptyset \\ B \in D^\Theta}} m(B) \tag{4}$$

In this paper, we will work with Shafer's model of the frame Θ, i.e. all elements θ_i of Θ are assumed truly exhaustive and exclusive (disjoint). Therefore $D^\Theta = 2^\Theta$ and the generalized belief functions reduces to classical ones. DSmT proposes a new efficient combination rules based on proportional conflict redistribution (PCR) principle for combining highly conflicting sources of evidence. Also, the classical pignistic transformation $BetP(.)$ [26] is replaced by the by the more effective $DSmP(.)$ transformation to estimate the subjective probabilities of hypotheses for classical decision-making. We just recall briefly the PCR fusion rule # 5 (PCR5) and Dezert-Smarandache Probabilistic (DSmP) transformation. All details, justifications with examples on PCR5 and DSmP can be found freely from the web in [22], Vols. 2 & 3 and will not be reported here.

• **The Proportional Conflict Redistribution Rule no. 5**: PCR5 is used generally to combine bba's in DSmT framework. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. Let $m_1(.)$ and $m_2(.)$ be two independent[2] bba's, then the PCR5 rule is defined as follows (see [22], Vol. 2 for full justification and examples): $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1)m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)}\right] \tag{5}$$

where all denominators in (5) are different from zero. If a denominator is zero, that fraction is discarded. Additional properties of PCR5 can be found in [9]. Extension of PCR5 for combining qualitative bba's can be found in [22], Vol. 2 & 3. All propositions/sets are in a canonical form. A variant of PCR5, called PCR6 has been proposed by Martin and Osswald in [22], Vol. 2, for combining $s > 2$ sources. The general formulas for PCR5 and PCR6 rules are given in [22], Vol. 2 also. PCR6 coincides with PCR5 for the fusion of two bba's.

• **DSmP probabilistic transformation**: $DSmP$ is a serious alternative to the classical pignistic transformation $BetP$ since it increases the probabilistic information content (PIC), i.e. it reduces Shannon entropy of the approximate subjective probability measure drawn from any bba – see [22], Vol. 3, Chap. 3 for details and the analytic expression of $DSmP_\epsilon(.)$. When $\epsilon > 0$ and when the masses of all singletons are zero, $DSmP_\epsilon(.) = BetP(.)$, where the well-known pignistic transformation $BetP(.)$ is defined by Smets in [26].

In the Evidential Reasoning framework, the decisions are usually achieved by computing the expected utilities of the acts using either the subjective/pignistic $BetP\{.\}$ (usually adopted in DST framework) or $DSmP(.)$ (as suggested in DSmT framework) as the probability function needed to compute expectations. Usually, one uses the maximum of the pignistic probability as decision criterion. The maximum of $BetP\{.\}$ is often considered as a balanced strategy between the two other strategies for decision making: the max of plausibility (optimistic strategy) or the max. of credibility (pessimistic strategy). The max of $DSmP(.)$ is considered as more efficient for practical applications since $DSmP(.)$ is more informative (it has a higher PIC value) than $BetP(.)$ transformation. The justification of DSmP as a fair and useful transformation for decision-making support can also be found in [10]. Note that in the binary frame case, all the aforementioned decision strategies yields same final decision.

III. Belief functions and MCDM

Two simple methods for MCDM under uncertainty are briefly presented: DSmT-AHP approach and Yager's OWA approach. The new Cautious OWA approach that we propose will be developed in the next section.

A. DSmT-AHP approach

DSmT-AHP aimed to perform a similar purpose as AHP [18], [19], SMART [30] or DS/AHP [1], [3], etc. that is to find the preferences rankings of the decision alternatives (DA), or groups of DA. DSmT-AHP approach consists in three steps:

- Step 1: we extend the construction of the matrix for taking into account the partial uncertainty (disjunctions) between

[2] i.e. each source provides its bba independently of the other sources.

possible alternatives. If no comparison is available between elements, then the corresponding elements in the matrix is zero. Each bba related to each (sub-) criterion is the normalized eigenvector associated with the largest eigenvalue of the "uncertain" knowledge matrix (as done in standard AHP approach).

- Step 2: we use the DSmT fusion rules, typically the PCR5 rule, to combine bba's drawn from step 1 to get a final MCDM priority ranking. This fusion step must take into account the different importances (if any) of criteria as it will be explained in the sequel.
- Step 3: decision-making can be based either on the maximum of belief, or on the maximum of the plausibility of DA, as well as on the maximum of the approximate subjective probability of DA obtained by different probabilistic transformations.

The MCDM problem deals with several criteria having different importances and the classical fusion rules cannot be applied directly as in step 2. In AHP, the fusion is done from the product of the bba's matrix with the weighting vector of criteria. Such AHP fusion is nothing but a simple component-wise weighted average of bba's and it doesn't actually process efficiently the conflicting information between the sources. It doesn't preserve the neutrality of a full ignorant source in the fusion. To palliate these problems, we have proposed a new solution for combining sources of different importances in [23]. Briefly, the reliability of a source is usually taken into account with Shafer's discounting method [21] defined by:

$$\begin{cases} m_\alpha(X) = \alpha \cdot m(X), \quad \text{for } X \neq \Theta \\ m_\alpha(\Theta) = \alpha \cdot m(\Theta) + (1-\alpha) \end{cases} \tag{6}$$

where $\alpha \in [0;1]$ is the reliability discounting factor. $\alpha = 1$ when the source is fully reliable and $\alpha = 0$ if the source is totally unreliable. We characterize the importance of a source by an importance factor β in $[0,1]$. β factor is usually not related with the reliability of the source and can be chosen to any value in $[0,1]$ by the designer for his/her own reason. By convention, $\beta = 1$ means the maximal importance of the source and $\beta = 0$ means no importance granted to this source. From this β factor, we define the importance discounting by

$$\begin{cases} m_\beta(X) = \beta \cdot m(X), \quad \text{for } X \neq \emptyset \\ m_\beta(\emptyset) = \beta \cdot m(\emptyset) + (1-\beta) \end{cases} \tag{7}$$

Here, we allow to deal with non-normal bba since $m_\beta(\emptyset) \geq 0$ as suggested by Smets in [24]. This new discounting preserves the specificity of the primary information since all focal elements are discounted with same importance factor. Based on this importance discounting, one can adapt PCR5 (or PCR6) rule for $N \geq 2$ discounted bba's $m_{\beta,i}(.)$, $i = 1,2,\ldots N$ to get with PCR5$_\emptyset$ fusion rule (see details in [23]) a resulting bba which is then normalized because in the AHP context, the importance factors correspond to the components of the normalized eigenvector $\mathbf{w}$. It is important to note that such importance discounting method cannot be used in DST when using Dempster-Shafer's rule of combination because this rule is not responding to the discounting of sources towards the empty set (see Theorem 1 in [23] for proof). The reliability and importance of sources can be taken into account easily in the fusion process and separately. The possibility to take them into account jointly is more difficult, because in general the reliability and importance discounting approaches do not commute, but when $\alpha_i = \beta_i = 1$. In order to deal both with reliabilities and importances factors and because of the non commutativity of these discountings, two methods have also been proposed in [23] and not reported here.

B. Yager's OWA approach

Let's introduce Yager's OWA approach [33] for decision making with belief structures. One considers a collection of q alternatives belonging to a set $A = \{A_1, A_2, \ldots, A_q\}$ and a finite set $S = \{S_1, S_2, \ldots, S_n\}$ of states of the nature. We assume that the payoff/gain C_ij of the decision maker in choosing A_i when S_j occurs are given by positive (or null) numbers. The payoffs $q \times n$ matrix is defined by $C = [C_{ij}]$ where $i = 1,\ldots,q$ and $j = 1,\ldots,n$ as in eq. (2). The decision-making problem consists in choosing the alternative $A^* \in A$ which maximizes the payoff to the decision maker given the knowledge on the state of the nature and the payoffs matrix C. $A^* \in A$ is called the best alternative or the solution (if any) of the decision-making problem. Depending the knowledge the decision-maker has on the states of the nature, he/she is face on different decision-making problems:
1 – Decision-making under certainty: only one state of the nature is known and certain to occur, say S_j. Then the decision-making solution consists in choosing $A^* = A_{i^*}$ with $i^* \triangleq \arg\max_i \{C_{ij}\}$.
2 – Decision-making under risk: the true state of the nature is unknown but one knows all the probabilities $p_j = P(S_j)$, $j = 1,\ldots,n$ of the possible states of the nature. In this case, we use the maximum of expected values for decision-making. For each alternative A_i, we compute its expected payoff $E[C_i] = \sum_j p_j \cdot C_{ij}$, then we choose $A^* = A_{i^*}$ with $i^* \triangleq \arg\max_i \{E[C_i]\}$.
3 – Decision-making under ignorance: one assumes no knowledge about the true state of the nature but that it belongs to S. In this case, Yager proposes to use the OWA operator assuming a given decision attitude taken by the decision-maker. Given a set of values/payoffs c_1, c_2, ..., c_n, OWA consists in choosing a normalized set of weighting factors $W = [w_1, w_2, \ldots w_n]$ where $w_j \in [0,1]$ and $\sum_j w_j = 1$ and for any set of values c_1, c_2, ..., c_n compute OWA$(c_1, c_2, \ldots, c_n)$ as

$$\text{OWA}(c_1, c_2, \ldots, c_n) = \sum_j w_j \cdot b_j \tag{8}$$

where b_j is the jth largest element in the collection c_1, c_2, ..., c_n. As seen in (8), the OWA operator is nothing but a simple weighted average of ordered values of a variable.
Based on such OWA operators, the idea consists for each alternative A_i, $i = 1,\ldots,q$ to choose a weighting vector $W_i = [w_{i1}, w_{i2}, \ldots w_{in}]$ and compute its OWA value $V_i \triangleq \text{OWA}(C_{i1}, C_{i2}, \ldots, C_{in}) = \sum_j w_{ij} \cdot b_{ij}$ where b_{ij} is the

jth largest element in the collection of payoffs C_{i1}, C_{i2},..., C_{in}. Then, as for decision-making under risk, we choose $A^* = A_{i^*}$ with $i^* \triangleq \arg\max_i\{V_i\}$. The determination of W_i depends on the decision attitude taken by the decision-maker. The pessimistic attitude considers for all $i = 1, 2, \ldots, q$, $W_i = [0, 0, \ldots, 0, 1]$. In this case, we assign to A_i the least payoff and we choose the best worst (the max of least payoffs). It is a Max-Min strategy since $i^* = \arg\max_i(\min_j C_{ij})$. The optimistic attitude considers for all $i = 1, 2, \ldots, q$, $W_i = [1, 0, \ldots, 0, 0]$. We commit to A_i its best payoff and we select the best best. It is a Max-Max strategy since $i^* = \arg\max_i(\max_j C_{ij})$. Between these two extreme attitudes, we can define an infinity of intermediate attitudes like the normative/neutral attitude (when or all $i = 1, 2, \ldots, q$, $W_i = [1/n, 1/n, \ldots, 1/n, 1/n]$) which corresponds to the simple arithmetic mean, or Hurwicz attitude (i.e. a weighted average of pessimistic and optimistic attitudes), etc. To justify the choice of OWA method, Yager defines an optimistic index $\alpha \in [0, 1]$ from the components of W_i and proposes to compute (by mathematical programming) the best weighting vector W_i corresponding to a priori chosen optimistic index and having the maximal entropy (dispersion). If $\alpha = 1$ (optimistic attitude) then of course $W_i = [1, 0, \ldots, 0, 0]$ and if $\alpha = 0$ (pessimistic attitude) then $W_i = [0, 0, \ldots, 0, 1]$. I theory, Yager's method doesn't exclude the possibility to adopt an hybrid attitude depending on the alternative we consider. In other words, we are not forced to consider the same weighting vectors for all alternatives.

Example 1: Let's take states $S = \{S_1, S_2, S_3, S_4\}$, alternatives $A = \{A_1, A_2, A_3\}$ and the payoffs matrix:

$$\begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & S_3 & S_4 \end{array} \\ \begin{pmatrix} 10 & 0 & 20 & 30 \\ 1 & 10 & 20 & 30 \\ 30 & 10 & 2 & 5 \end{pmatrix} \end{array} \tag{9}$$

If one adopts the pessimistic attitude in choosing $W_1 = W_2 = W_3 = [0, 0, 0, 1]$, then one gets for each alternative A_i, $i = 1, 2, 3$ the following values of OWA's: $V_1 = \text{OWA}(10, 0, 20, 30) = 0$, $V_2 = \text{OWA}(1, 10, 20, 30) = 1$ and $V_3 = \text{OWA}(30, 10, 2, 5) = 2$. The final decision will be the alternative V_3 since it offers the best expected payoff.

If one adopts the optimistic attitude in choosing $W_1 = W_2 = W_3 = [1, 0, 0, 0]$, then one gets for each alternative A_i, $i = 1, 2, 3$ the following values of OWA's: $V_1 = \text{OWA}(10, 0, 20, 30) = 30$, $V_2 = \text{OWA}(1, 10, 20, 30) = 30$ and $V_3 = \text{OWA}(30, 10, 2, 5) = 30$. All alternatives offer the same expected payoff and thus the final decision must be chosen randomly or purely ad-hoc since there is no best alternative.

If one adopts the normative attitude in choosing $W_1 = W_2 = W_3 = [1/4, 1/4, 1/4, 1/4]$ (i.e. one assumes that all states of nature are equiprobable), then one gets: $V_1 = \text{OWA}(10, 0, 20, 30) = 60/4$, $V_2 = \text{OWA}(1, 10, 20, 30) = 61/4$ and $V_3 = \text{OWA}(30, 10, 2, 5) = 47/4$. The final decision will be the alternative V_2 since it offers the best expected payoff.

4 – Decision-making under uncertainty: this corresponds to the general case where the knowledge on the states of the nature is characterized by a belief structure. Clearly, one assumes that a priori knowledge on the frame S of the different states of the nature is given by a bba $m(.) : 2^S \to [0, 1]$. This case includes all previous cases depending on the choice of $m(.)$. Decision under certainty is characterized by $m(S_j) = 1$; Decision under risk is characterized by $m(s) > 0$ for some states $s \in S$; Decision under full ignorance is characterized by $m(S_1 \cup S_2 \cup \ldots \cup S_n) = 1$, etc. Yager's OWA for decision-making under uncertainty combines the schemes used for decision making under risk and ignorance. It is based on the derivation of a *generalized expected value* C_i of payoff for each alternative A_i as follows:

$$C_i = \sum_{k=1}^{r} m(X_k) V_{ik} \tag{10}$$

where r is the number of focal elements of the belief structure $(S, m(.))$. $m(X_k)$ is the mass of belief of the focal element $X_k \in 2^S$, and V_{ik} is the payoff we get when we select A_i and the state of the nature lies in X_k. The derivation of V_{ik} is done similarly as for the decision making under ignorance when restricting the states of the nature to the subset of states belonging to X_k only. Therefore for A_i and a focal element X_k, instead of using all payoffs C_{ij}, we consider only the payoffs in the set $M_{ik} = \{C_{ij} | S_j \in X_k\}$ and $V_{ik} = \text{OWA}(M_{ik})$ for some decision-making attitude chosen a priori. Once generalized expected values C_i, $i = 1, 2, \ldots, q$ are computed, we select the alternative which has its highest C_i as the best alternative (i.e. the final decision). The principle of this method is very simple, but its implementation can be quite greedy in computational resources specially if one wants to adopt a particular attitude for a given level of optimism, specially if the dimension of the frame S is large: one needs to compute by mathematical programming the weighting vectors generating the optimism level having the maximum of entropy. As illustrative example, we take Yager's example[3] [33] with a pessimistic, optimistic and normative attitudes.

Example 2: Let's take states $S = \{S_1, S_2, S_3, S_4, S_5\}$ with associated bba $m(S_1 \cup S_3 \cup S_4) = 0.6$, $m(S_2 \cup S_5) = 0.3$ and $m(S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5) = 0.1$. Let's also consider alternatives $A = \{A_1, A_2, A_3, A_4\}$ and the payoffs matrix:

$$C = \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} \tag{11}$$

The $r = 3$ focal elements of $m(.)$ are $X_1 = S_1 \cup S_3 \cup S_4$, $X_2 = S_2 \cup S_5$ and $X_3 = S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5$. X_1 and X_2 are partial ignorances and X_3 is the full ignorance. One considers the following submatrix (called bags by Yager) for

[3]There is a mistake/typo error in original Yager's example [33].

the derivation of V_{ik}, for $i = 1, 2, 3, 4$ and $k = 1, 2, 3$.

$$M(X_1) = \begin{bmatrix} M_{11} \\ M_{21} \\ M_{31} \\ M_{41} \end{bmatrix} = \begin{bmatrix} 7 & 12 & 13 \\ 12 & 5 & 11 \\ 9 & 3 & 10 \\ 6 & 11 & 15 \end{bmatrix}$$

$$M(X_2) = \begin{bmatrix} M_{12} \\ M_{22} \\ M_{32} \\ M_{42} \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 10 & 2 \\ 13 & 9 \\ 9 & 4 \end{bmatrix}$$

$$M(X_3) = \begin{bmatrix} M_{13} \\ M_{23} \\ M_{33} \\ M_{43} \end{bmatrix} = \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} = C$$

• Using pessimistic attitude, and applying the OWA operator on each row of $M(X_k)$ for $k = 1$ to r, one gets finally[4]: $V(X_1) = [V_{11}, V_{21}, V_{31}, V_{41}]^t = [7, 5, 3, 6]^t$, $V(X_2) = [V_{12}, V_{22}, V_{32}, V_{42}]^t = [5, 2, 9, 4]^t$ and $V(X_3). = [V_{13}, V_{23}, V_{33}, V_{43}]^t = [5, 2, 3, 4]^t$. Applying formula (10) for $i = 1, 2, 3, 4$ one gets finally the following generalized expected values using vectorial notation:

$$[C_1, C_2, C_3, C_4]^t = \sum_{k=1}^{r=3} m(X_k) \cdot V(X_k) = [6.2, 3.8, 4.8, 5.2]^t$$

According to these values, the best alternative to take is A_1 since it has the highest generalized expected payoff.

• Using optimistic attitude, one takes the max value of each row, and applying OWA on each row of $M(X_k)$ for $k = 1$ to r, one gets: $V(X_1) = [V_{11}, V_{21}, V_{31}, V_{41}]^t = [13, 12, 10, 15]^t$, $V(X_2) = [V_{12}, V_{22}, V_{32}, V_{42}]^t = [6, 10, 13, 9]^t$, and $V(X_3) = [V_{13}, V_{23}, V_{33}, V_{43}]^t = [13, 12, 13, 15]^t$. One finally gets $[C_1, C_2, C_3, C_4]^t = [10.9, 11.4, 11.2, 13.2]^t$ and the best alternative to take with optimistic attitude is A_4 since it has the highest generalized expected payoff.

• Using normative attitude, one takes $W_1 = W_2 = W_3 = W_4 = [1/|X_k|, 1/|X_k|, \ldots, 1/|X_k|]$ where $|X_k|$ is the cardinality of the focal element X_k under consideration. The number of elements in W_i is equal to $|X_k|$. The generalized expected values are $[C_1, C_2, C_3, C_4]^t = [9.1, 8.3, 8.4, 9.4]^t$ and the best alternative with the normative attitude is A_4 (same as with optimistic attitude) since it has the highest generalized expected payoff.

C. Using expected utility theory

In this section, we propose to use a much simpler approach than OWA Yager's approach for decision making under uncertainty. The idea is to approximate the bba $m(.)$ by a subjective probability measure through a given probabilistic transformation. We suggest to use either $BetP$ or (better) $DSmP$ transformations for doing this as explained in [22] (Vol.3, Chap. 3). Let's take back the previous example and compute the $BetP(.)$ and $DSmP_\epsilon(.)$ values from $m(.)$. One gets the same values in this particular example for any $\epsilon > 0$ because we don't have singletons as focal elements of $m(.)$, which is normal. Here $BetP(S_1) = DSmP(S_1) = 0.22$, $BetP(S_2) = DSmP(S_2) = 0.17$, $BetP(S_3) = DSmP(S_3) = 0.22$, $BetP(S_4) = DSmP(S_4) = 0.22$ and $BetP(S_5) = DSmP(S_2) = 0.17$. Based on these probabilities, we can compute the expected payoffs for each alternative as for decision making under risk (e.g. for C_1, we get $7 \cdot 0.22 + 5 \cdot 0.17 + 12 \cdot 0.22 + 13 \cdot 0.22 + 6 \cdot 0.17 = 8.91$). For the 4 alternatives, we finally get:

$$E_{BetP}[C] = E_{DSmP}[C] = [8.91, 8.20, 8.58, 9.25]^t$$

According to these values, one sees that the best alternative with this pignistic or DSm attitude is A_4 (same as with Yager's optimistic or normative attitudes) since it offers the highest pignistic or DSm expected payoff. This much simpler approach must be used with care however because there is a loss of information through the approximation of the bba $m(.)$ into any subjective probability measure. Therefore, we do not recommend to use it in general.

IV. The new COWA-ER approach

Yager's OWA approach is based on the choice of given attitude measured by an optimistic index in $[0, 1]$ to get the weighting vector W. How is chosen such an index/attitude ? This choice is ad-hoc and very disputable for users. What to do if we don't know which attitude to adopt ? The rational answer to this question is to consider the results of the two extreme attitudes (pessimistic and optimistic ones) jointly and try to develop a new method for decision under uncertainty based on the imprecise valuation of alternatives. This is the approach developed in this paper and we call it Cautious OWA with Evidential Reasoning (COWA-ER) because it adopts the cautious attitude (based on the possible extreme attitudes) and ER, as explained in the sequel.

Let's take back the previous example and take the pessimistic and optimistic valuations of the expected payoffs. The expected payoffs $E[C_i]$ are imprecise since they belong to interval $[C_i^{\min}, C_i^{\max}]$ where bounds are computed with extreme pessimistic and optimistic attitudes, and one has

$$E[C] = \begin{bmatrix} E[C_1] \\ E[C_2] \\ E[C_3] \\ E[C_4] \end{bmatrix} \subset \begin{bmatrix} [6.2; 10.9] \\ [3.8; 11.4] \\ [4.8; 11.2] \\ [5.2; 13.2] \end{bmatrix}$$

Therefore, one has 4 sources of information about the parameter associated with the best alternative to choose. For decision making under imprecision, we propose to use here again the belief functions framework and to adopt the following very simple COWA-ER methodology based on the following four steps:

- Step 1: normalization of imprecise values in $[0, 1]$;
- Step 2: conversion of each normalized imprecise value into elementary bba $m_i(.)$;
- Step 3: fusion of bba $m_i(.)$ with an efficient combination rule (typically PCR5);

[4] where X^t denotes the transpose of X.

- Step 4: choice of the final decision based on the resulting combined bba.

Let's describe in details each step of COWA-ER. In step 1, we divide each bound of intervals by the max of the bounds to get a new normalized imprecise expected payoff vector $E^{Imp}[C]$. In our example, one gets:

$$E^{Imp}[C] = \begin{bmatrix} [6.2/13.2; 10.9/13.2] \\ [3.8/13.2; 11.4/13.2] \\ [4.8/13.2; 11.2/13.2] \\ [5.2/13.2; 13.2/13.2] \end{bmatrix} \approx \begin{bmatrix} [0.47; 0.82] \\ [0.29; 0.86] \\ [0.36; 0.85] \\ [0.39; 1.00] \end{bmatrix}$$

In step 2, we convert each imprecise value into its bba according to a very natural and simple transformation [7]. Here, we need to consider as frame of discernment, the finite set of alternatives $\Theta = \{A_1, A_2, A_3, A_4\}$ and the sources of belief associated with them obtained from the normalized imprecise expected payoff vector $E^{Imp}[C]$. The modeling for computing a bba associated to the hypothesis A_i from any imprecise value $[a; b] \subseteq [0; 1]$ is very simple and is done as follows:

$$\begin{cases} m_i(A_i) = a, \\ m_i(\bar{A}_i) = 1 - b \\ m_i(A_i \cup \bar{A}_i) = m_i(\Theta) = b - a \end{cases} \quad (12)$$

where $\bar{A}_i$ is the complement of A_i in Θ. With such simple conversion, one sees that $Bel(A_i) = a$, $Pl(A_i) = b$. The uncertainty is represented by the length of the interval $[a; b]$ and it corresponds to the imprecision of the variable (here the expected payoff) on which is defined the belief function for A_i. In the example, one gets:

Alternatives A_i	$m_i(A_i)$	$m_i(\bar{A}_i)$	$m_i(A_i \cup \bar{A}_i)$
A_1	0.47	0.18	0.35
A_2	0.29	0.14	0.57
A_3	0.36	0.15	0.49
A_4	0.39	0	0.61

Table I
BASIC BELIEF ASSIGNMENTS OF THE ALTERNATIVES

In step 3, we need to combine bba's $m_i(.)$ by an efficient rule of combination. Here, we suggest to use the PCR5 rule proposed in DSmT framework since it has been proved very efficient to deal with possibly highly conflicting sources of evidence. PCR5 has been already applied successfully in all applications where it has been used so far [22]. We call this COWA-ER method based on PCR5 as COWA-PCR5. Obviously, we could replace PCR5 rule by any other rule (DS rule, Dubois& Prade, Yager's rule, etc and thus define easily COWA-DS, COWA-DP, COWA-Y, etc variants of COWA-ER. This is not the purpose of this paper and this has no fundamental interest in this presentation. The result of the combination of bba's with PCR5 for our example is given in of Table II.

The last step 4 is the decision-making from the resulting bba of the fusion step 3. This problem is recurrent in the theory of belief functions and several attitudes are also possible as explained at the end of section II. Table III shows what are the values of credibilities, plausibilities, $BetP$ and $DSmP_{\epsilon=0}$ for each alternative in our example.

Focal Element	$m_{PCR5}(.)$
A_1	0.2488
A_2	0.1142
A_3	0.1600
A_4	0.1865
$A_1 \cup A_4$	0.0045
$A_2 \cup A_4$	0.0094
$A_1 \cup A_2 \cup A_4$	0.0236
$A_3 \cup A_4$	0.0075
$A_1 \cup A_3 \cup A_4$	0.0198
$A_2 \cup A_3 \cup A_4$	0.0374
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.1883

Table II
FUSION OF THE FOUR ELEMENTARY BBA'S WITH PCR5

A_i	$Bel(A_i)$	$BetP(A_i)$	$DSmP(A_i)$	$Pl(A_i)$
A_1	0.2488	0.3126	0.3364	0.4850
A_2	0.1142	0.1863	0.1623	0.3729
A_3	0.1600	0.2299	0.2242	0.4130
A_4	0.1865	0.2712	0.2771	0.4521

Table III
CREDIBITITY AND PLAUSIBILITY OF A_i

Based on the results of Table III, it is interesting to note that, in this example, there is no ambiguity in the decision making whatever the attitude is taken by the decision-maker (the max of Bel, the max of Pl, the max of BetP or the max of DSmP), the decision to take will always be A_1. Such behavior is probably not general in all problems, but at least it shows that in some cases like in Yager's example, the ambiguity in decision can be removed when using COWA-PCR5 instead of OWA which is an advantage of our approach. It is worth to note that Shannon entropy of BetP is $H_{BetP} = 1.9742$ bits is bigger than Shannon entropy of DSmP is $H_{DSmP} = 1.9512$ bits which is normal since DSmP has been developed for increasing the PIC value.

Advantages and extension of COWA-ER: COWA-PCR5 allows also to take easily a decision, not only on a single alternative, but also if one wants on a group/subset of alternatives satisfying a min of credibility (or plausibility level) selected by the decision-maker. Using such approach, it is of course very easy to discount each bba $m_i(.)$ entering in the fusion process using reliability or importance discounting techniques which makes this approach more appealing and flexible for the user than classical OWA. COWA-PCR5 is simpler to implement because it doesn't require the evaluation of all weighting vectors for the bags by mathematical programming. Only extreme and very simple weighting vectors $[1, 0, \ldots, 0]$ and $[0, \ldots, 0, 1]$ are used in COWA-ER. Of course, COWA-ER can also be extended directly for the fusion of several sources of informations when each source can provide a payoffs matrix. It suffices to apply COWA-ER on each matrix to get the bba's of step 3, then combine them with PCR5 (or any other rule) and then apply step 4 of COWA-ER. We can also discount each

source easily if needed. All these advantages makes COWA-ER approach very flexible and appealing for MCDM under uncertainty. In summary, the original OWA approach considers several alternatives A_i evaluated in the context of different uncertain scenarii and includes several ways (pessimistic, optimistic, hurwicz, normative) to interpret and aggregate the evaluations with respect to a given scenario. COWA-ER uses simultaneously the two extreme pessimistic and optimistic decision attitudes combined with an efficient fusion rule as shown on Figure 3. In order to save computational resources (if required), we also have proposed a less efficient OWA approach using the classical concept of expected utility based on DSmP or BetP.

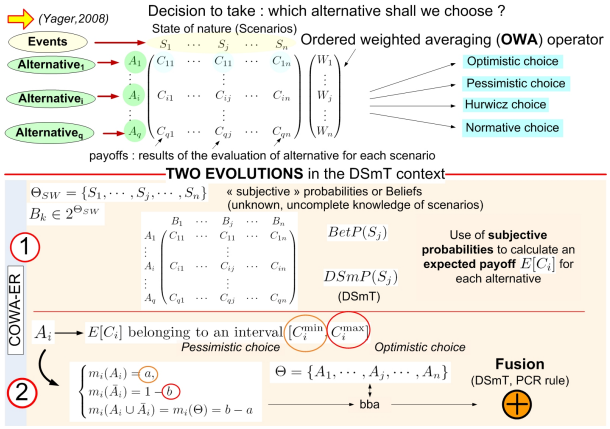


Figure 3. COWA-ER: Two evolutions of Yager's OWA method.

V. Conclusion

In this work, Yager's Ordered Weighted Averaging (OWA) operators are extended and simplified with evidential reasoning (ER) for MCDM under uncertainty. The new Cautious OWA-ER method is very flexible and requires less computational load than classical OWA. COWA-ER improves the existing framework for MCDM since it can deal also with several more or less reliable sources. Further developments are now planned to combine uncertainty about states of the world with the imperfection and uncertainty of alternatives evaluations as previously introduced in the ER-MCDA and DSmT-AHP methods in order to connect them with COWA-ER.

References

[1] M. Beynon, B. Curry, P.H. Morgan, *The Dempster-Shafer theory of evidence: An alternative approach to multicriteria decision modelling*, Omega, Vol. 28, No. 1, pp. 37–50, 2000.

[2] M. Beynon, D. Cosker, D. Marshall, *An expert system for multi-criteria decision making using Dempster-Shafer theory*, Expert Syst. with Appl. Vol. 20, No. 4, pp. 357–367, 2001.

[3] M. Beynon, *DS/AHP method: A mathematical analysis, including an understanding of uncertainty*, Eur. J. of Oper. Research, Vol. 140, pp. 148–164, 2002.

[4] M. Beynon, *Understanding local ignorance and non-specificity within the DS/AHP method of multi-criteria decision making*, Eur. J. of Oper. Research, Vol. 163, pp. 403–417, 2005.

[5] M. Beynon, *A method of aggregation in DS/AHP for group decision-making with non-equivalent importance of individuals in the group*, Comp. and Oper. Research, No. 32, pp. 1881–1896, 2005.

[6] D. Bouyssou, *Modelling inaccurate determination, uncertainty, imprecision using multiple criteria*, Lecture Notes in Econ. & Math. Syst., 335:78–87, 1989.

[7] J. Dezert,*Autonomous navigation with uncertain reference points using the PDAF*, in Multitarget-Multisensor Tracking, Vol 2, pp 271–324, Y. Bar-Shalom Editor, Artech House, 1991.

[8] J. Dezert, J.-M. Tacnet, M. Batton-Hubert, F. Smarandache,*Multi-criteria decision making based on DSmT-AHP* in Proc. of Belief 2010 Int. Workshop, Brest, France, 1-2 April, 2010.

[9] J. Dezert, F. Smarandache, *Non Bayesian conditioning and deconditioning*, in Proc. of Belief 2010 Int. Workshop, Brest, France, 2010.

[10] D. Han, J. Dezert, C. Han, Y. Yang, *Is Entropy Enough to Evaluate the Probability Transformation Approach of Belief Function?*, in Proceedings of Fusion 2010 conference, Edinburgh, UK, July 2010.

[11] Z. Hua, B. Gong, X. Xu, *A DS-AHP approach for multi-attribute decision making problem with incomplete information*, Expert Systems with Appl., 34(3):2221–2227, 2008.

[12] D. Kahneman, A. Tversky, *Prospect theory : An analysis of decision under risk*, Econometrica, 47:263–291, 1979.

[13] A. Martin, A.-L. Jousselme, C. Osswald, *Conflict measure for the discounting operation on belief functions*, Proc. of Fusion 2008 Int. Conf.

[14] P. Mongin, *Expected utility theory*, Handbook. of Economic Methodology, pp. 342–350, Edward Elgar, London, 1997.

[15] M. S. Ozdemir,T. L. Saaty, *The unknown in decision making: What to do about it ?* Eur. J.of Oper. Research, 174(1):349–359, 2006.

[16] B. Roy, *Main sources of inaccurate determination, uncertainty and imprecision in decision models*, Math. & Comput. Modelling, 12 (10-11):1245–1254, 1989.

[17] B. Roy, *Paradigms and challenges*, in Multiple Criteria Decision Analysis : State of the art surveys, Vol.78 of Int. Series in Oper. Research and& Management Sci. (Chap. 1), pp. 1–24, Springer, 2005.

[18] T.L. Saaty, *A scaling method for priorities in hierarchical structures*, J. of Math. Psych., Vol. 15, PP. 59–62, 1977.

[19] T.L. Saaty, *The Analytical Hierarchy Process*, McGraw Hill, 1980.

[20] T.L. Saaty, *Fundamentals of decision making and priority theory with the analytic hierarchy process*, Vol. VI of the AHP series, RWL Publ., Pittsburgh, PA, USA,2000.

[21] G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, 1976.

[22] F. Smarandache, J. Dezert (Editors), *Advances and Applications of DSmT for Information Fusion*, American Research Press, Rehoboth, Vol.1-3, 2004-2009 - see http://fs.gallup.unm.edu//DSmT.htm.

[23] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, in Proc. of Fusion 2010 Int. Conf., Edinburgh, UK, July 2010.

[24] P. Smets, *The Combination of Evidence in the Transferable Belief Model*, IEEE Trans. PAMI 12, pp. 447–458, 1990.

[25] P. Smets, R. Kennes, *The transferable belief model*, Artif. Intel., 66(2), pp. 191-234, 1994.

[26] Ph. Smets, *Decision making in the TBM: the necessity of the pignistic transformation*, Int. J. of Approx. Reas., Vol. 38, pp. 133–147, 2005.

[27] T. J. Stewart, *Dealing with uncertainties in MCDA*, in Multiple Criteria Decision Analysis: State of the art surveys, Vol.78 of Int. Series in Op. Res. & Manag. Sci. (chapter 11), pp. 445–466, Springer, 2005.

[28] J.-M. Tacnet, M. Batton-Hubert, J. Dezert, *Information fusion for natural hazards in mountains* in [22], Vol. 3, 2009.

[29] J.-M. Tacnet, M. Batton-Hubert, J. Dezert, *A two-step fusion process for multi-criteria decision applied to natural hazards in mountains*, Proc. of Belief 2010 Int. Workshop, Brest, France, 2010.

[30] D. Von Winterfeldt, W. Edwards *Decision analysis and behavioral research*, Cambridge Univ. Press, 1986.

[31] R. Yager, *On ordered weighted averaging operators in multi-criteria decision making*, EEE Trans. on SMC, 18:183–190, 1988

[32] R. Yager *Induced ordered weighted averaging operators*, IEEE Trans. on SMC, Part B: Cybernetics, Vol. 29, No. 2, pp:141–150, April, 1999.

[33] R. Yager, *Decision making under Dempster-Shafer uncertainties*, Studies in Fuzziness and Soft Computing, 219:619–632, 2008.

[34] L. Zadeh, *On the validity of Dempster's rule of combination*, Memo M79/24, Univ. of California, Berkeley, USA, 1979.

Sigmoidal Model for Belief Function-Based Electre Tri Method

Jean Dezert
Jean-Marc Tacnet

Originally published as Dezert J., Tacnet J.-M., S*igmoidal Model for Belief Function-based Electre Tri method*, in Belief 2012, Compiègne, May 2012, and reprinted with permission.

Abstract. Main decision-making problems can be described into choice, ranking or sorting of a set of alternatives or solutions. The principle of Electre TRI (ET) method is to sort alternatives a_i according to criteria g_j into categories C_h whose lower and upper limits are respectively b_h and b_{h+1}. The sorting procedure is based on the evaluations of outranking relations based firstly on calculation of partial concordance and discordance indexes and secondly on global concordance and credibility indexes. In this paper, we propose to replace the calculation of the original concordance and discordance indexes of ET method by a more effective sigmoidal model. Such model is part of a new Belief Function ET (BF-ET) method under development and allows a comprehensive, elegant and continuous mathematical representation of degree of concordance, discordance and the uncertainty level which is not directly taken into account explicitly in the classical Electre Tri.

1 Introduction

The Electre Tri (ET) method, developed by Yu [13], remains one of the most successful and applied methods for multiple criteria decision aiding (MCDA) sorting problems [5]. ET method assigns a set of given alternatives $a_i \in \mathbf{A}$, $i = 1,2,\ldots,n$ according to criteria g_j, $j = 1,2,\ldots,m$ to a pre-defined (and ordered) set of categories $C_h \in \mathbf{C}$, $h = 1,2,\ldots,p+1$ whose lower and upper limits are respectively b_h and b_{h+1} for all $h = 1,\ldots,p$), with $b_0 \leq b_1 \leq b_2 \leq \ldots \leq b_{h-1} \leq b_h \leq \ldots \leq b_p$. The assignment of an alternative a_i to a category C_h (limited by profiles b_h and $b_h + 1$) consists in four steps involving at first the computation of global concordance $c(a_i,b_h)$ and discordance $d(a_i,b_h)$ indexes[1] (steps 1 & 2), secondly their fusion into a credibility

[1] Themselves computed from partial concordance and discordance indexes based on a given set criteria $g_j(.)$, $j \in \mathbf{J}$.

index $\rho(a_i,b_h)$ (step 3), and finally the decision and choice of the category based on the evaluations of outranking relations [13, 6] (step 4). The partial concordance index $c_j(a_i,b_h)$ measures the concordance of a_i and b_h in the assertion *"a_i is at least as good as b_h"*. The partial discordance index $d_j(a_i,b_h)$ measures the opposition of a_i and b_h in the assertion *"a_i is at least as good as b_h"*. The global concordance index $c(a_i,b_h)$ measures the concordance of a_i and b_h on all criteria in the assertion *"a_i outranks b_h"*. The degree of credibility of the outranking relation denoted as $\rho(a_i,b_h)$ expresses to which extent *"a_i outranks b_h"* according to $c(a_i,b_h)$ and $d_j(a_i,b_h)$ for all criteria. The main steps of ET method are described below:

1. **Concordance Index:** The concordance index $c(a_i,b_h) \in [0,1]$ between the alternative a_i and the category C_h is computed as the weighted average of partial concordance indexes $c_j(a_i,b_h)$, that is

$$c(a_i,b_h) = \sum_{j\in\mathbf{J}} w_j c_j(a_i,b_h) \tag{1}$$

where the weights $w_i \in [0,1]$ represent the relative importance of each criterion $g_j(.)$ in the evaluation of the global concordance index. They must satisfy $\sum_{j\in\mathbf{J}} w_j = 1$. The partial concordance index $c_j(a_i,b_h) \in [0,1]$ based on a given criterion $g_j(.)$ is computed from the difference of the criteria evaluated for the profil b_h, and the criterion evaluated for the alternative a_i. If the difference $g_j(b_h) - g_j(a_i)$ is less (or equal) to a given preference threshold $q_j(g_j(b_h))$ then a_i and C_h are considered as different based on the criterion $g_j(.)$ so that a preference of a_i with respect to C_h can be clearly done. If the difference $g_j(b_h) - g_j(a_i)$ is strictly greater to another given threshold $p_j(g_j(b_h))$ then a_i and C_h are considered as indifferent (similar) based on $g_j(.)$). When $g_j(b_h) - g_j(a_i) \in [q_j(g_j(b_h)), p_j(g_j(b_h))]$, the partial concordance index $c_j(a_i,b_h)$ is computed from a linear interpolation. Mathematically, the partial concordance index is obtained by:

$$c_j(a_i,b_h) \triangleq \begin{cases} 1 & \text{if} \quad g_j(b_h) - g_j(a_i) \leq q_j(g_j(b_h)) \\ 0 & \text{if} \quad g_j(b_h) - g_j(a_i) > p_j(g_j(b_h)) \\ \frac{g_j(a_i) + p_j(g_j(b_h)) - g_j(b_h)}{p_j(g_j(b_h)) - q_j(g_j(b_h))} & \text{otherwise} \end{cases} \tag{2}$$

2. **Discordance Index:** The discordance index between the alternative a_i and the category C_h depends on a possible veto condition expressed by the choice of a veto threshold $v_j(g_j(b_h))$ imposed on some criterion $g_j(.)$. The (global) discordance index $d(a_i,b_h)$ is computed from the partial discordance indexes:

$$d_j(a_i,b_h) \triangleq \begin{cases} 1 & \text{if} \quad g_j(b_h) - g_j(a_i) > v_j(g_j(b_h)) \\ 0 & \text{if} \quad g_j(b_h) - g_j(a_i) \leq p_j(g_j(b_h)) \\ \frac{g_j(b_h) - g_j(a_i) - p_j(g_j(b_h))}{v_j(g_j(b_h)) - p_j(g_j(b_h))} & \text{otherwise} \end{cases} \tag{3}$$

One defines by $\mathbf{V}$ the set of indexes $j \in \mathbf{J}$ where the veto applies (where the partial discordance index is greater than the global concordance index), that is

$$\mathbf{V} \triangleq \{j \in \mathbf{J} | d_j(a_i, b_h) > c(a_i, b_h)\} \tag{4}$$

Then a global discordance index can be defined [12] as

$$d(a_i, b_h) \triangleq \begin{cases} 1 & \text{if} \quad \mathbf{V} = \emptyset \\ \prod_{j \in \mathbf{V}} \frac{1 - d_j(a_i, b_h)}{1 - c_j(a_i, b_h)} & \text{if} \quad \mathbf{V} \neq \emptyset \end{cases} \tag{5}$$

3. **Global Credibility Index:** In ET method, the (global) credibility index $\rho(a_i, b_h)$ is computed by the simple discounting of the concordance index $c(a_i, b_h)$ given by (1) by the discordance index (discounting factor) $d(a_i, b_h)$ given in (5). Mathematically, this is given by

$$\rho(a_i, b_h) = c(a_i, b_h) d(a_i, b_h) \tag{6}$$

4. **Assignment Procedure:** The assignment of a given action a_i to a certain category C_h results from the comparison of a_i to the profile defining the lower and upper limits of the categories. For a given category limit b_h, this comparison relies on the credibility of the assertions a_i outranks b_h. Once all credibility indexes $\rho(a_i, b_h)$ for $i = 1, 2, \ldots, m$ and $h = 1, 2, \ldots, k$ have been computed, the assignment matrix $\mathbf{M} \triangleq [\rho(a_i, b_h)]$ is available for helping in the final decision-making process. In ELECTRE TRI method, a simple λ-cutting level strategy (for a given choice of $\lambda \in [0.5, 1]$) is used in order to transform the fuzzy outranking relation into a crisp one to determine if each alternative outranks (or not) each category. This is done by testing if $\rho(a_i, b_h) \geq \lambda$. If the inequality is satisfied, it means that indeed a_i outranks the category C_h. Based on outranking relations between all pairs of alternatives and profiles of categories, two approches are proposed in ELECTRE TRI to finally assign the alternatives into categories, see [5] for details:

   - Pessimistic (conjunctive) approach: a_i is compared with b_k, b_{k-1}, b_{k-2}, ..., until a_i outranks b_h where $h \leq k$. The alternative a_i is then assigned to the highest category C_h if $\rho(a_i, b_h) \geq \lambda$ for a given threshold λ.
   - Optimistic (disjunctive) approach: a_i is compared with b_1, b_2, ... b_h, ... until b_h outranks a_i. The alternative a_i is assigned to the lowest category C_h for which the upper profile b_h is preferred to a_i.

The objective and motivation of this paper is to develop a new Belief Function based ET method taking into account the potential of BF to model uncertainties. The whole BF-ET method is under development and will be presented and evaluated on a detailed practical example in a forthcoming publication. Due to space limitation constraints, we just present here what we propose to compute the new concordance and discordance indexes useful in our BF-ET.

2 Limitations of the Classical Electre Tri

ET method remains rather based on heuristic approach than on a theoretical one for each of its steps. Belief functions can improve ET method because of their ability to model and manage conflicting as well as uncertainty information in a theoretical framework. We only focus here on steps 1 and 2 and we propose a solution to overcome their limitations in the next section.

Example 1: Let's consider $g_j(a_i) \in [0,100]$, and let's take $g_j(b_h) = 50$ and the following thresholds: $q_j(g_j(b_h)) = 20$ (indifference threshold), $p_j(g_j(b_h)) = 25$ (preference threshold) and $v_j(g_j(b_h) = 40$ (veto threshold). Then the local concordance and discordances indexes obtained in steps 1 and 2 of ET are shown on the Fig. 1.

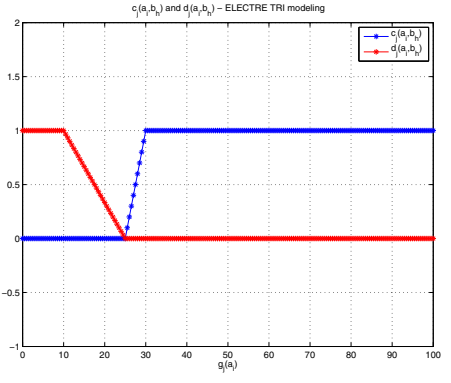


Fig. 1 Example of partial concordance and discordance indexes.

From this very simple example, one sees that ET modeling of partial concordance and discordance indexes is not very satisfactory since there is no clear (explicit and consistent) modeling of the uncertainty area where the action a_i is not totally discordant, nor totally concordant with the profile b_h. In such simplistic modeling, there exist points $g_j(a_i)$ (lying on the slope of the blue or red curves) that can be not totally concordant while being totally not discordant (and vice-versa), which is counter-intuitive and rather abnormal. This drawback will be solved using our new sigmoidal basic belief assignment (bba) modeling presented in the next section.

3 Sigmoidal Model for Concordance and Discordance Indexes

In fact, there are several ways to compute partial concordances and discordances indexes and to combine them in order to provide the global credibility indexes $\rho(a_i, b_h)$. Electre Tri proposes a simple and basic approach based on hard thresholding techniques for doing this. It can fail to work efficiently in practice in some cases, or may require a lot of experience to calibrate/tune all setting parameters in order to apply it to get pertinent results for decision-making support. Usually, a sensitivity analysis must be done very carefully before applying ET in real applications. Here,

we propose a more flexible approach based on sigmoidal modeling where no hard thresholding technique is required.

In ET approach, we are mainly concerned in the evaluation of the credibility indexes $\rho(a_i,b_h)\in[0,1]$ for $i=1,2,\ldots,m$ and $h=1,2,\ldots,k$ (step 3) from which the final decision (assignment) will be drawn in step 4. Step 3 is conditioned by the results of steps 1 and 2 which can be improved using belief functions. For such purpose, we consider, a binary frame of discernment[2] $\Theta\triangleq\{c,\bar{c}\}$ where c means that the alternative a_i is concordant with the assertion "a_i is at least as good as profile b_h", and $\bar{c}$ means that the alternative a_i is opposed (discordant) to this assertion. This must obviously be done with all the assertions to check in the ET framework. The basic idea is for each pair (a_i,b_h) to evaluate its bba $m_{ih}(.)$ defined on the power-set of Θ, denoted 2^Θ. Such bba's have of course to be defined from the combination (fusion) of the local bba's $m^j_{ih}(.)$ evaluated from each possible criteria $g_j(.)$ (as in steps 1 and 2). The main issue is to derive the local bba's $m^j_{ih}(.)$ defined in 2^Θ from the knowledge of the criteria $g_j(.)$ and preference, indifference and veto thresholds $p_j(g_j(b_h))$, $q_j(g_j(b_h))$ and $v_j(g_j(b_h))$ respectively. It turns out that this can be easily obtained from the new method of construction of bba presented in [4] and adapted here in the ET context as follows:

• Let $g_j(a_i)$ be the evaluation of the criterion $g_j(.)$ for the alternative a_i, following ET approach when $g_j(a_i)\geq g_j(b_h)-q_j(g_j(b_h))$ then the belief in concordance c must be high (close to one), whereas it must be low (close to zero) as soon as $g_j(a_i)<g_j(b_h)-p_j(g_j(b_h))$. Similarly, the belief in discordance $\bar{c}$ must be high (close to one) if $g_j(a_i)<g_j(b_h)-v_j(g_j(b_h))$, and it must be low (close to zero) when $g_j(a_i)\geq g_j(b_h)-p_j(g_j(b_h))$. Such behavior can be modeled directly from the sigmoid functions defined by $f_{s,t}(g)\triangleq 1/(1+e^{-s(g-t)})$ where g is the criterion magnitude of the alternative under consideration; t is the abscissa of the inflection point of the sigmoid. $s/4$ is the slope[3] of the tangent at the inflection point. It can be easily verified that the bba $m^j_{ih}(.)$ satisfying the expected behavior can be obtained by the fusion[4] of the two following simple bba's defined by: where the abscisses of inflection points are given by $t_c=g_j(b_h)-\frac{1}{2}(p_j(g_j(b_h))+q_j(g_j(b_h)))$ and $t_{\bar{c}}=g_j(b_h)-\frac{1}{2}(p_j(g_j(b_h))+v_j(g_j(b_h)))$ and the parameters s_c and $s_{\bar{c}}$ are given by[5] $s_c=4/(p_j(g_j(b_h))-q_j(g_j(b_h)))$ and $s_{\bar{c}}=4/(v_j(g_j(b_h))-p_j(g_j(b_h)))$.

Table 1 Construction of $m_1(.)$ and $m_2(.)$.

focal element	$m_1(.)$	$m_2(.)$
c	$f_{s_c,t_c}(g)$	0
$\bar{c}$	0	$f_{-s_{\bar{c}},t_{\bar{c}}}(g)$
$c\cup\bar{c}$	$1-f_{s_c,t_c}(g)$	$1-f_{-s_{\bar{c}},t_{\bar{c}}}(g)$

[2] Here we assume that Shafer's model holds, that is $c\cap\bar{c}=\emptyset$.

[3] i.e. the ratio of the vertical and horizontal distances between two points on a line; zero if the line is horizontal, undefined if it is vertical.

[4] With averaging rule, PCR5 rule, or Dempster-Shafer rule [8].

[5] The coefficient 4 appearing in s_c and $s_{\bar{c}}$ expressions comes from the fact that for a sigmoid of parameter s, the tangent at its inflection point is $s/4$.

• From the setting of threshold parameters $p_j(g_j(b_h))$, $q_j(g_j(b_h))$ and $v_j(g_j(b_h))$, it is easy to compute the parameters of the sigmoids (t_c, s_c) and $(t_{\bar{c}}, t_{\bar{c}})$, and thus to get the values of bba's $m_1(.)$ and $m_2(.)$. Once this has been done the local bba $m^j_{ih}(.)$ is computed by the fusion (denoted $\oplus$) of bba's $m_1(.)$ and $m_2(.)$, that is $m^j_{ih}(.) = [m_1 \oplus m_2](.)$. As shown in [4], the choice of a particular rule of combination (Dempster, PCR5, or hybrid rule) has only a little impact on the result of the combined bba $m^j_{ih}(.)$. But since PCR5 proposes a better management of conflicting bba's yielding to more specific results than with other rules [1], we use it to combine $m_1(.)$ with $m_2(.)$ to compute $m^j_{ih}(.)$ associated with the criterion $g_j(.)$ and the pair (a_i, b_h). In adopting such sigmoidal modeling, we get now from $m^j_{ih}(.)$ a fully consistent and elegant representation of local concordance $c_j(a_i, b_h)$ (step 1 of ET), local discordance $d_j(a_i, b_h)$ (step 2 of ET), as well as of the local uncertainty $u_j(a_i, b_h)$ by considering: $c_j(a_i, b_h) \triangleq m^j_{ih}(c) \in [0,1]$, $d_j(a_i, b_h) \triangleq m^j_{ih}(\bar{c}) \in [0,1]$ and $u_j(a_i, b_h) \triangleq m^j_{ih}(c \cup \bar{c}) \in [0,1]$. Of course, one has also $c_j(a_i, b_h) + d_j(a_i, b_h) + u_j(a_i, b_h) = 1$.

4 Example of a Sigmoidal Model

If one takes back the example 1, the inflection points of the sigmoids $f_1(g) \triangleq f_{s_c,t_c}(g)$ and $f_2(g) \triangleq f_{-s_{\bar{c}},t_{\bar{c}}}(g)$ have the following abscisses $t_c = 50 - (25+20)/2 = 27.5$ and $t_{\bar{c}} = 50 - (25+40)/2 = 17.5$ and parameters $s_c = 4/(25-20) = 4/5 = 0.8$ and $s_{\bar{c}} = 4/(40-25) = 4/15 \approx 0.2666$. The two sigmoids $f_1(g_j(a_i))$ and $f_2(g_j(a_i))$ are shown on the Fig. 2.

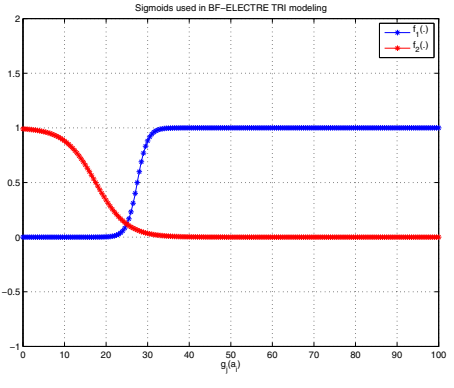


Fig. 2 $f_1(g_j(a_i))$ and $f_2(g_j(a_i))$ sigmoids.

It is interesting to note the resemblance of Fig. 2 with Fig. 1. From these sigmoids, the bba's $m_1(.)$ and $m_2(.)$ are computed according to Table 1 and shown on the Figure 3.

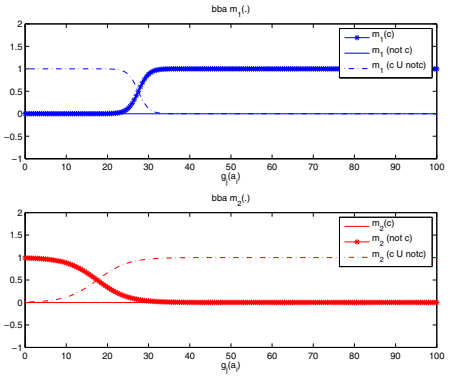


Fig. 3 Bba's $m_1(.)$ and $m_2(.)$ to combine.

The construction of the consistent bba $m^j_{ih}(.)$ is obtained by the PCR5 fusion of the bba's $m_1(.)$ and $m_2(.)$. The result is shown on Fig. 4.

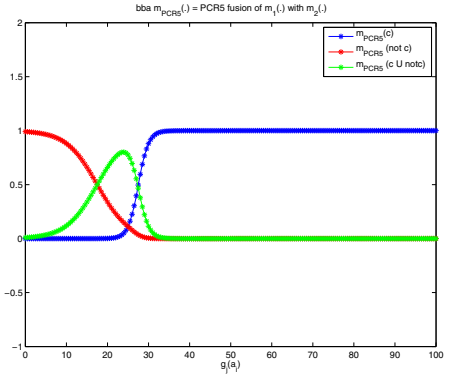


Fig. 4 $m^j_{ih}(.)$ obtained from the PCR5 fusion of $m_1(.)$ with $m_2(.)$.

From this new sigmoidal modeling, we can compute the local bba's $m^j_{ih}(.)$ derived from the knowledge of criterion $g_j(.)$ and setting parameters. This is a smooth appealing and elegant technique to build all the local bba's: no hard thresholding is necessary because of the continuity of sigmoid functions.

One can then compute the global concordance and discordance indexes of steps 1 and 2 from the computation of the combined bba $m_{ih}(.)$ resulting of the fusion of local bba's $m^j_{ih}(.)$ taking eventually into account their importance and reliability[6] (if one wants). This can be done using the recent fusion techniques proposed in [9], or by a simple weighted averaging. From $m_{ih}(.)$ we can use the same credibility index as in step 3 of ET, or just skip this third step and define a decision-making based directly on the bba $m_{ih}(.)$ using classical approaches used in belief function framework (say the max of belief, plausibility, or pignistic probability, etc).

[6] In classical ET, the reliability of criteria is not taken into account.

5 Conclusions

After a brief presentation of the classical ET method, we have proposed a new approach to model and compute the concordance and discordance indexes based on belief functions in order to overcome the limitations of steps 1 and 2 of the ET approach. The advantages of our modeling is to provide an elegant and simple way not only to compute the concordance and discordance indexes, but also the uncertainty level that may occur when information appears partially concordant and discordant. The Improvements of other steps of ET method are under development. In future reaserch works, we will evaluate and compare on real MCDA problem our BF-ET with the original ET method and with other belief functions based methods already available in MCDA frameworks [10, 11].

References

1. Dezert, J., Smarandache, F.: Proportional Conflict Redistribution Rules for Information Fusion. In: [8], vol. 2, pp. 3–68 (2006)
2. Dezert, J., Smarandache, F.: An introduction to DSmT. In: [8], vol. 3, pp. 3–73 (2009)
3. Dezert, J., Smarandache, F., Tacnet, J.-M., Batton-Hubert, M.: Multi-criteria decision making based on DSmT/AHP. In: International Workshop on Belief Functions, Brest, France (April 2010)
4. Dezert, J., Liu, Z., Mercier, G.: Edge Detection in Color Images Based on DSmT. In: Proceedings of Fusion 2011, Chicago, USA (July 2011)
5. Figueira, J., Mousseau, V., Roy, B.: ELECTRE methods. In: Multiple Criteria Decision Analysis: State of Art Surveys, ch. 4. Springer Science+Business Media Inc. (2005)
6. Mousseau, V., Slowinski, R., Zielniewicz, P.: Electre tri 2.0a - methological guide and user's manual - document no 111. In: Cahier et Documents du Lamsade, Lamsade, Université Paris-Dauphine, Paris (1999)
7. Shafer, G.: A mathematical theory of evidence. Princeton University Press (1976)
8. Smarandache, F., Dezert, J.: Advances and applications of DSmT for information fusion (Collected works), vol. 1-3. American Research Press (2004-2009), `http://fs.gallup.unm.edu/DSmT.htm`
9. Smarandache, F., Dezert, J., Tacnet, J.-M.: Fusion of sources of evidence with different importances and reliabilities. In: Proc. of Fusion 2010 Conf., Edinburgh, UK (July 2010)
10. Tacnet, J.-M., Dezert, J.: Cautious OWA and Evidential Reasoning for Decision Making under Uncertainty. In: Proceedings of Fusion 2011, Chicago, USA (July 2011)
11. Tacnet, J.-M., Batton-Hubert, M., Dezert, J.: A two-step fusion process for multi-criteria decision applied to natural hazards in mountains. In: Proceedings of International Workshop on the Theory of Belief Functions, Belief 2010, Brest (France), April 1-2 (2010)
12. Tervonen, T., Figueira, J.R., Lahdelma, R., Dias, J.A., Salminen, P.: A stochastic method for robustness analysis in sorting problems. European Journal of Operational Research 192, 236–242 (2009)
13. Yu, W.: Aide multicritère à la décision dans le cadre de la problématique du tri: Concepts, méthodes et applications. Ph.D Thesis, University Paris-Dauphine, France (1992)

Hierarchical Proportional Redistribution Principle for Uncertainty Reduction and BBA Approximation*

Jean Dezert
Deqiang Han
Zhun-ga Liu
Jean-Marc Tacnet

Originally published as Dezert J., Han D., Liu Z., Tacnet J.-M., *Hierarchical Proportional Redistribution principle for uncertainty reduction and bba approximation*, WCICA2012 Beijing, China, July 2012, and reprinted with permission.

***Abstract*—Dempster-Shafer evidence theory is very important in the fields of information fusion and decision making. However, it always brings high computational cost when the frames of discernments to deal with become large. To reduce the heavy computational load involved in many rules of combinations, the approximation of a general belief function is needed. In this paper we present a new general principle for uncertainty reduction based on hierarchical proportional redistribution (HPR) method which allows to approximate any general basic belief assignment (bba) at a given level of non-specificity, up to the ultimate level 1 corresponding to a Bayesian bba. The level of non-specificity can be adjusted by the users. Some experiments are provided to illustrate our proposed HPR method.**

***Index Terms*—Belief functions, hierarchical proportional redistribution (HPR), evidence combination, belief approximation.**

I. Introduction

Dempster-Shafer evidence theory, also called belief function theory [1], is an interesting and flexible tool to deal with imprecision and uncertainty for approximate reasoning. It has been widely used in many applications, e.g., information fusion, pattern recognition and decision making [2].

Although evidence theory is successful in uncertainty modeling and reasoning, high computational cost is a drawback which is often raised against evidence theory [2]. In fact, the computational cost of evidence combination increases exponentially with respect to the size of the frame of discernment (FOD) [3]–[5]. To resolve such a problem, three major types of approaches have been proposed by he researchers.

The first type is to propose efficient procedures for performing exact computations. For example, Kennes [6] proposed an optimal algorithm for Dempster's rule of combination. Barnett's work [7] and other works [8] are also the representatives.

The second type is composed of Monte-Carlo techniques. See details in the paper of Moral and Salmeron [9].

The third type is to approximate (or simplify) a belief function to a simpler one. The papers of Voorbraak [4], Dubois and Prade [10] are seminal works in this type of approaches. Tessem proposed the famous $k-l-x$ approximation approach [3]. Grabisch proposed some approaches [11], which can build a bridge between belief functions and other types of uncertainty measures or functions, e.g., probabilities, possibilities and k-additive belief function (those belief functions whose cardinality of the focal elements are at most of k). Based on pignistic transformation in transferable belief model (TBM), Burger and Cuzzolin proposed two types of k-additive belief functions [12]. Denœux uses hierarchical clustering strategy to implement the inner and outer approximation of belief functions [13].

In this paper, we focus on the approximation approach of belief functions. This first reason obviously is that it can reduce the computational cost of evidence combination. Furthermore, human find that it is not intuitive to attach meaning to focal elements with large cardinality [14]. Belief approximation can either reduce the number or the cardinalities of focal elements, or both of them can be reduced. Thus by using belief function approximation, we can obtain a representation which is more intuitive and easier to process. We propose a new method called hierarchical proportional redistribution (HPR), which is a general principle for uncertainty reduction, to approximate any general basic belief assignment (bba) at a given level of non-specificity, up to the ultimate level 1 corresponding to a Bayesian bba. That is, our proposed approach can generate an intermediate object between probabilities and original belief function. The level of non-specificity can be controlled by the users through the adjusting of maximum cardinality of remaining focal element. Our proposed approach can be considered as a generalized k-additive belief approximation. Some experiments are provided to illustrate our proposed HPR approach and to compare it with other approximation approaches.

II. Basics in evidence theory

In Dempster-Shafer evidence theory [1], the frame of discernment (FOD) denoted by Θ is a basic concept. The elements in Θ are mutually exclusive. Suppose that 2^Θ denotes the powerset of FOD and define the function $m : 2^\Theta \rightarrow [0, 1]$ as the basic belief assignment (bba) satisfying:

$$\sum\nolimits_{A\subseteq\Theta} m(A) = 1,\ m(\emptyset) = 0 \qquad (1)$$

*This work was supported by National Natural Science Foundation of China (Grant No.61104214), Fundamental Research Funds for the Central Universities, China Postdoctoral Science Foundation (No.20100481337, No.201104670)and Research Fund of Shaanxi Key Laboratory of Electronic Information System Integration (No.201101Y17).

A bba is also called a mass function. Belief function (Bel) and plausibility function (Pl) are defined below, respectively:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{2}$$

$$pl(A) = \sum_{A \cap B \neq \emptyset} m(B) \tag{3}$$

Suppose that $m_1, m_2, ..., m_n$ are n mass functions, Dempster's rule of combination is defined in (4):

$$m(A) = \begin{cases} 0, & A = \emptyset \\ \frac{\sum_{\cap A_i = A} \prod_{1 \leq i \leq n} m_i(A_i)}{\sum_{\cap A_i \neq \emptyset} \prod_{1 \leq i \leq n} m_i(A_i)}, & A \neq \emptyset \end{cases} \tag{4}$$

Dempster's rule of combination is used in DST to implement the combination of bodies of evidence (BOEs).

Evidence theory has been widely used in many application fields due to its capability of approximate reasoning and processing of uncertain information. However, as referred in introduction section, there also exists the drawback of high computational cost in evidence combination. Several approaches have been proposed accordingly, which includes efficient algorithms [6]–[8] for evidence combination, the Monte-Carlo techniques and the approach of belief function approximation [9]. We prefer to use the belief approximation approach [10]–[13] to reduce the computational cost needed in the combination operation because the approximation approach reduces the computational cost and also allow to deal with smaller-size focal elements, which is more intuitive for human to catch the meaning [14]. In the next section, we recall some well-known basic approximation approaches.

III. BBA APPROXIMATION APPROACHES

1) $k - l - x$ approach: The approach of $k - l - x$ was proposed by Tessem [3]. The simplified or compact bba obtained by using $k - l - x$ satisfies:

1) keep no less than k focal elements;
2) keep no more than l focal elements;
3) the mass assignment to be deleted is no greater than x.

In algorithm of $k-l-x$, the focal elements of a original bba are sorted according to their mass assignments. Such algorithm chooses the first p focal elements such that $k \leq p \leq l$ and such that the sum of the mass assignments of these first p focal elements is no less than $1 - x$. The deleted mass assignments are redistributed to the other focal elements through a normalization.

2) k-additive belief function approximation: Given a bba $m : 2^\Theta \rightarrow [0, 1]$, the k-additive belief function [11], [12] induced by the mass assignment is defined in Eq.(5). Suppose that $B \subseteq \Theta$,

$$\begin{cases} m_k(B) = m(B) + \sum_{A \supset B, A \subseteq \Theta, |A| > k} \frac{m(A) \cdot |B|}{\mathcal{N}(|A|,k)}, & \forall\, |B| \leq k \\ m_k(B) = 0, & \forall\, |B| > k \end{cases} \tag{5}$$

where

$$\mathcal{N}(|A|, k) = \sum_{j=1}^{k} \binom{|A|}{j} \cdot j = \sum_{j=1}^{k} \frac{|A|!}{(j-1)!(|A|-j)!} \tag{6}$$

is average cardinality of the subsets of A of size at most k.

It can be seen that for k-additive belief approximation, the maximum cardinality of available focal elements is no greater than k.

In this section, $k - l - x$ approach and k-additive belief function approximation approach are introduced, which will be compared with our proposed approach introduced in next section. There also other types of bba approximation approximation approaches, see details in related references.

IV. HIERARCHICAL PROPORTIONAL REDISTRIBUTION APPROXIMATION

In this paper we propose a hierarchical bba approximation approach called hierarchical proportional redistribution (HPR), which provides a new way to reduce step by step the mass committed to uncertainties. Ultimately an approximate measure of subjective probability can be obtained if needed, i.e. a so-called Bayesian bba in [1]. It must be noticed that this procedure can be stopped at any step in the process and thus it allows to reduce the number of focal elements in a given bba in a consistent manner to diminish the size of the core of a bba and thus reduce the complexity (if needed) when applying also some complex rules of combinations. We present here the general principle of hierarchical and proportional reduction of uncertainties in order to obtain approximate bba's at different non-specificity level we expect. The principle of redistribution of uncertainty to more specific elements of the core at any given step of the process follows the proportional redistribution already proposed in the (non hierarchical) DSmP transformation proposed recently in [5]. Thus the proposed HPR can be considered as a bba approximation approach inspired by the idea of DSmP.

Let's first introduce two new notations for convenience and for concision:

1) Any element of cardinality $1 \leq k \leq n$ of the power set 2^Θ will be denoted, by convention, by the generic notation $X(k)$. For example, if $\Theta = \{A, B, C\}$, then $X(2)$ can denote the following partial uncertainties $A \cup B$, $A \cup C$ or $B \cup C$, and $X(3)$ denotes the total uncertainty $A \cup B \cup C$.
2) The proportional redistribution factor (ratio) of width n involving elements X and Y of the powerset is defined as (for $X \neq \emptyset$ and $Y \neq \emptyset$)

$$R_s(Y, X) \triangleq \frac{m(Y) + \epsilon \cdot |X|}{\sum_{\substack{Y \subset X \\ |X|-|Y|=s}} m(Y) + \epsilon \cdot |X|} \tag{7}$$

where ϵ is a small positive number introduced here to deal with particular cases where $\sum_{\substack{Y \subset X \\ |X|-|Y|=s}} m(Y) = 0$. By convention, we will denote $R(Y, X) \triangleq R_1(Y, X)$ when we use the proportional redistribution factors of width $s = 1$.

The HPR is obtained by a step by step (recursive) proportional redistribution of the mass $m(X(k))$ of a given uncertainty $X(k)$ (partial or total) of cardinality $2 \leq k \leq n$ to all the least

specific elements of cardinality $k-1$, i.e. to all possible $X(k-1)$, until $k=2$ is reached. The proportional redistribution is done from the masses of belief committed to $X(k-1)$ as done classically in DSmP transformation. The "hierarchical" masses $m_h(.)$ are recursively (backward) computed as follows. Here $m_{h(k)}$ represents the approximate bba obtained at the step $n-k$ of HPR, i.e., it has the maximum focal element cardinality of k.

$$m_{h(n-1)}(X(n-1)) = m(X(n-1)) + \sum_{\substack{X(n)\supset X(n-1)\\ X(n),X(n-1)\in 2^\Theta}} [m(X(n))\cdot R(X(n-1),X(n))];$$

$$m_{h(n-1)}(A) = m(A), \forall |A| < n-1 \tag{8}$$

$m_{h(n-1)}(\cdot)$ is the bba obtained at the first step of HPR ($n-(n-1)=1$), the maximum focal element cardinality of $m_{h(n-1)}$ is $n-1$.

$$m_{h(n-2)}(X(n-2)) = m(X(n-2)) + \sum_{\substack{X(n-1)\supset X(n-2)\\ X(n-2),X(n-1)\in 2^\Theta}} [m_{h(n-1)}(X(n-1))\cdot R(X(n-2),X(n-1))]$$

$$m_{h(n-2)}(A) = m_{h(n-1)}(A), \forall |A| < n-2 \tag{9}$$

$m_{h(n-2)}(\cdot)$ is the bba obtained at the second step of HPR ($n-(n-2)=2$), the maximum focal element cardinality of $m_{h(n-2)}$ is $n-2$.

This hierarchical proportional redistribution process can be applied similarly (if one wants) to compute $m_{h(n-3)}(.)$, $m_{h(n-4)}(.)$, ..., $m_{h(2)}(\cdot)$, $m_{h(1)}(\cdot)$ with

$$m_{h(2)}(X(2)) = m(X(2)) + \sum_{\substack{X(3)\supset X(2)\\ X(3),X(2)\in 2^\Theta}} [m_{h(3)}(X(3))\cdot R(X(2),X(3))]$$

$$m_{h(2)}(A) = m_{h(3)}(A), \forall |A| < n-2 \tag{10}$$

$m_{h(2)}(\cdot)$ is the bba obtained at the first step of HPR ($n-2$), the maximum focal element cardinality of $m_{h(2)}$ is 2.

Mathematically, for any $X(1)\in\Theta$, i.e. any $\theta_i\in\Theta$ a Bayesian belief function can be obtained by HPR approach in deriving all possible steps of proportional redistributions of partial ignorances in order to get

$$m_{h(1)}(X(1)) = m(X(1)) + \sum_{\substack{X(2)\supset X(1)\\ X(1),X(2)\in 2^\Theta}} [m_{h(2)}(X(2))\cdot R(X(1),X(2))] \tag{11}$$

In fact, $m_{h(1)}(\cdot)$ is a probability transformation, called here the Hierarchical DSmP (HDSmP). Since $X(n)$ is unique and corresponds only to the full ignorance $\theta_1\cup\theta_2\cup\ldots\cup\theta_n$, the expression of $m_h(X(n-1))$ in Eq.(10) just simplifies as

$$m_{h(n-1)}(X(n-1)) = m_h(X(n-1)) + m(X(n))\cdot R(X(n-1),X(n)) \tag{12}$$

Because of the full proportional redistribution of the masses of uncertainties to the elements least specific involved in these uncertainties, no mass of belief is lost during the step by step hierarchical process and thus at any step of HPR, the sum of masses of belief is kept to one, and of course the Hierarchial DSmP also satisfies $\sum_{X(1)\in 2^\Theta} m_h(X(1)) = 1$.

Remark: For some reasons depending of applications, it is also possible to easily modify this HPR approach with little effort into a constrained HPR version (CHPR for short) by forcing the masses of some partial ignorances of cardinality $k+1$ to be (proportionally) redistributed back only to a subset of the partial ignorances of cardinality k included in them. This possibility has not be detailed here due to space limitation constraint and its little technical interest.

V. Examples

In this section we show in details how HPR can be applied on very simple different examples. So let's examine the three following examples based on a simple 3D frame of discernment $\Theta=\{\theta_1,\theta_2,\theta_3\}$ satisfying Shafer's model.

A. Example 1

Let's consider the following bba:

$$m(\theta_1)=0.10,\quad m(\theta_2)=0.17,\quad m(\theta_3)=0.03,$$
$$m(\theta_1\cup\theta_2)=0.15,\quad m(\theta_1\cup\theta_3)=0.20,$$
$$m(\theta_2\cup\theta_3)=0.05,\quad m(\theta_1\cup\theta_2\cup\theta_3)=0.30.$$

We apply the hierarchical proportional redistribution (HPR) principle with $\epsilon=0$ in this example because there is no mass of belief equal to zero. It can be verified that the result obtained with small positive ϵ parameter remains (as expected) numerically very close to that obtained with $\epsilon=0$.

The first step of HPR consists in redistributing back $m(\theta_1\cup\theta_2\cup\theta_3)=0.30$ committed to the full ignorance to the elements $\theta_1\cup\theta_2$, $\theta_1\cup\theta_3$ and $\theta_2\cup\theta_3$ only, because these elements are the only elements of cardinality 2 that are included in $\theta_1\cup\theta_2\cup\theta_3$. Applying the Eq. (8) with $n=3$, one gets when $X(2)=\theta_1\cup\theta_2$, $\theta_1\cup\theta_3$ and $\theta_1\cup\theta_2$ the following masses.

$$m_{h(2)}(\theta_1\cup\theta_2) = m(\theta_1\cup\theta_2) + m(X(3))\cdot R(\theta_1\cup\theta_2, X(3)) = 0.15 + (0.30\cdot 0.375) = 0.2625$$

because $R(\theta_1\cup\theta_2, X(3)) = \frac{0.15}{0.15+0.20+0.05} = 0.375$.
Similarly, one gets

$$m_{h(2)}(\theta_1\cup\theta_3) = m(\theta_1\cup\theta_3) + m(X(3))\cdot R(\theta_1\cup\theta_3, X(3)) = 0.20 + (0.30\cdot 0.5) = 0.35$$

because $R(\theta_1 \cup \theta_3, X(3)) = \frac{0.20}{0.15+0.20+0.05} = 0.5$, and also

$$\begin{aligned} m_{h(2)}(\theta_2 \cup \theta_3) &= m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3)) \\ &= 0.05 + (0.30 \cdot 0.125) = 0.0875 \end{aligned}$$

because $R(\theta_2 \cup \theta_3, X(3)) = \frac{0.05}{0.15+0.20+0.05} = 0.125$.

Now, we go to the next step of HPR principle and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use Eq. (11) for doing this as follows:

$$\begin{aligned} m_{h(1)}(\theta_1) &= m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3) \\ &\approx 0.10 + (0.2625 \cdot 0.3703) + (0.35 \cdot 0.7692) \\ &= 0.10 + 0.0972 + 0.2692 = 0.4664 \end{aligned}$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.10}{0.10 + 0.17} \approx 0.3703$$

$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.10}{0.10 + 0.03} \approx 0.7692$$

Similarly, one gets

$$\begin{aligned} m_{h(1)}(\theta_2) &= m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3) \\ &\approx 0.10 + (0.2625 \cdot 0.6297) + (0.0875 \cdot 0.85) \\ &= 0.17 + 0.1653 + 0.0744 = 0.4097 \end{aligned}$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17}{0.10 + 0.17} \approx 0.6297$$

$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17}{0.17 + 0.03} = 0.85$$

and also

$$\begin{aligned} m_{h(1)}(\theta_3) &= m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3) \\ &\approx 0.03 + (0.35 \cdot 0.2307) + (0.0875 \cdot 0.15) \\ &= 0.03 + 0.0808 + 0.0131 = 0.1239 \end{aligned}$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.03}{0.10 + 0.03} \approx 0.2307$$

$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.03}{0.17 + 0.03} = 0.15$$

Hence, the result of final step of HPR is:

$$\begin{aligned} &m_{h(1)}(\theta_1) = 0.4664, \quad m_{h(1)}(\theta_2) = 0.4097, \\ &m_{h(1)}(\theta_3) = 0.1239. \end{aligned}$$

We can easily verify that

$$m_{h(1)}(\theta_1) + m_{h(1)}(\theta_2) + m_{h(1)}(\theta_3) = 1.$$

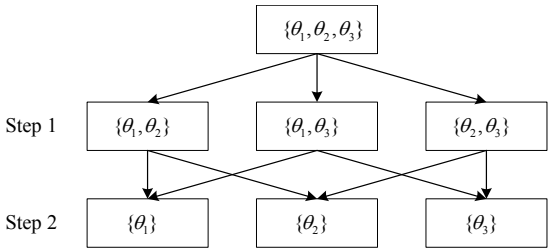


Figure 1. Illustration of Example 1

Table I
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 1.

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k=3$	$k=2$	$k=1$
θ_1	0.1000	0.1000	0.4664
θ_2	0.1700	0.1700	0.4097
θ_3	0.0300	0.0300	0.1239
$\theta_1 \cup \theta_2$	0.1500	0.2625	0.0000
$\theta_1 \cup \theta_3$	0.2000	0.3500	0.0000
$\theta_2 \cup \theta_3$	0.0500	0.0875	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

The procedure can be illustrated in Fig. 1 below. The approximate bba at each step with different maximum focal elements' cardinality are listed in Table I.

To compare our proposed HPR with the approach of $k-l-x$, we set the parameters of $k-l-x$ to obtain bba's with equal focal element number with HPR at each step. In Example 1, for HPR at first step, it can obtain a bba with 6 focal elements. Thus we set $k=l=6, x=0.4$ for $k-l-x$ to obtain a bba with 6 focal elements. Similarly, for HPR at second step, it can obtain a bba with 3 focal elements. Thus we set $k=l=3, x=0.4$ for $k-l-x$. Based on the approach of $k-l-x$, the results are in Table II.

Table II
EXPERIMENTAL RESULTS OF $k-l-x$ FOR EXAMPLE 1

Focal elements	$m(\cdot)$ obtained by $k-l-x$	
	$k=l=6$	$k=l=3$
θ_1	0.1031	0.0000
θ_2	0.1753	0.2573
θ_3	0.0000	0.0000
$\theta_1 \cup \theta_2$	0.1546	0.0000
$\theta_1 \cup \theta_3$	0.2062	0.2985
$\theta_2 \cup \theta_3$	0.0515	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3093	0.4478

B. Example 2

Let's consider the following bba:

$$\begin{aligned} &m(\theta_1) = 0, \quad m(\theta_2) = 0.17, \quad m(\theta_3) = 0.13, \\ &m(\theta_1 \cup \theta_2) = 0.20, \quad m(\theta_1 \cup \theta_3) = 0.20, \\ &m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30 \end{aligned}$$

The first step of HPR consists in redistributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, and $\theta_1 \cup \theta_3$ only, because these elements are the only elements of cardinality 2 that are included in $\theta_1 \cup \theta_2 \cup \theta_3$.

Applying Eq. (8) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses

$$\begin{aligned} m_{h(2)}(\theta_1 \cup \theta_2) &= m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3)) \\ &= 0.20 + (0.3 \cdot 0.5) = 0.35 \end{aligned}$$

because

$$R(\theta_1 \cup \theta_2, X(3)) = \frac{0.20}{0.20 + 0.20 + 0.00} = 0.5$$

Similarly, one gets

$$\begin{aligned} m_{h(2)}(\theta_1 \cup \theta_3) &= m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3)) \\ &= 0.20 + (0.3 \cdot 0.5) = 0.35 \end{aligned}$$

because

$$R(\theta_1 \cup \theta_3, X(3)) = \frac{0.20}{0.20 + 0.20 + 0.00} = 0.5$$

and also

$$\begin{aligned} m_{h(2)}(\theta_2 \cup \theta_3) &= m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3)) \\ &= 0.00 + (0.3 \cdot 0.0) = 0.0 \end{aligned}$$

because

$$R(\theta_2 \cup \theta_3, X(3)) = \frac{0.0}{0.20 + 0.20 + 0.00} = 0$$

Now, we go to the next step of HPR principle and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use Eq. (11) for doing this as follows:

$$\begin{aligned} m_{h(1)}(\theta_1) &= m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3) \\ &\approx 0.00 + (0.35 \cdot 0.00) + (0.35 \cdot 0.00) \\ &= 0.00 + 0.00 + 0.00 = 0.00 \end{aligned}$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.00}{0.00 + 0.17} = 0.00$$

$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.00}{0.00 + 0.13} = 0.00$$

Similarly, one gets

$$\begin{aligned} m_{h(1)}(\theta_2) &= m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3) \\ &\approx 0.17 + (0.35 \cdot 1) + (0.00 \cdot 0.5667) \\ &= 0.17 + 0.35 + 0.00 = 0.52 \end{aligned}$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17}{0.00 + 0.17} = 1$$

$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17}{0.17 + 0.13} \approx 0.5667$$

and also

$$\begin{aligned} m_{h(1)}(\theta_3) &= m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3) \\ &\approx 0.13 + (0.35 \cdot 1) + (0.00 \cdot 0.4333) \\ &= 0.13 + 0.35 + 0.00 = 0.48 \end{aligned}$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.13}{0.13 + 0.00} = 1$$

$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.13}{0.17 + 0.13} \approx 0.4333$$

Hence, the result of final step of HPR is

$$m_{h(1)}(\theta_1) = 0.00, \quad m_{h(1)}(\theta_2) = 0.52, \quad m_{h(1)}(\theta_3) = 0.48$$

and we can easily verify that

$$m_{h(1)}(\theta_1) + m_{h(1)}(\theta_2) + m_{h(1)}(\theta_3) = 1.$$

The HPR procedure of Example 2 with $\epsilon = 0$ is Fig. 2.

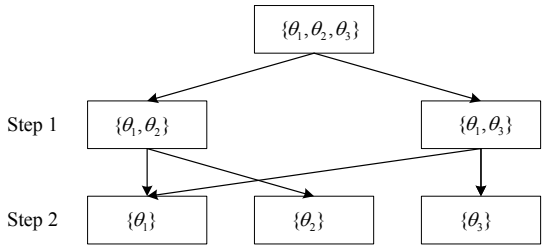


Figure 2. Illustration of Example 2.

If one takes $\epsilon = 0$, there is no mass that will be reassigned to $\{\theta_2 \cup \theta_3\}$ as illustrated in Fig. 2. But if one takes $\epsilon > 0$, HPR procedure of Example 2 is the same as that illustrated in Fig. 1, i.e., there also exist masses redistributed to $\{\theta_2 \cup \theta_3\}$ as illustrated in Fig. 1. That's the difference between Fig. 1 and Fig. 2.

Suppose that $\epsilon = 0.001$, the HPR calculation procedure is as follows.

The first step of HPR consists in distributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$. Applying the Eq. (8) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses

$$\begin{aligned} m_{h(2)}(\theta_1 \cup \theta_2) &= m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3)) \\ &= 0.20 + (0.3 \cdot 0.4963) = 0.3489 \end{aligned}$$

because

$$\begin{aligned} R(\theta_1 \cup \theta_2, X(3)) &= \frac{0.20 + 0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)} \\ &= 0.4963 \end{aligned}$$

$$\begin{aligned} m_{h(2)}(\theta_1 \cup \theta_3) &= m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3)) \\ &= 0.20 + (0.3 \cdot 0.4963) = 0.3489 \end{aligned}$$

because

$$R(\theta_1 \cup \theta_2, X(3)) = \frac{0.20 + 0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)} = 0.4963$$

$$m_{h(2)}(\theta_2 \cup \theta_3) = m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3)) = 0.00 + (0.3 \cdot 0.0073) = 0.0022$$

because

$$R(\theta_2 \cup \theta_3, X(3)) = \frac{0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)} = 0.0073$$

Now, we go to the next step of HPR principle and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use Eq. (11) for doing this as follows:

$$\begin{aligned} m_{h(1)}(\theta_1) &= m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3) \\ &\approx 0.00 + (0.3489 \cdot 0.0115) + (0.3489 \cdot 0.0149) \\ &= 0.00 + 0.0040 + 0.0052 = 0.0092 \end{aligned}$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.00 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.17 + 0.001 \cdot 2)} = 0.0115$$

$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.00 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)} = 0.0149$$

Similarly, one gets

$$\begin{aligned} m_{h(1)}(\theta_2) &= m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3) \\ &\approx 0.17 + (0.3489 \cdot 0.9885) + (0.0022 \cdot 0.5658) \\ &= 0.17 + 0.3449 + 0.0012 = 0.5161 \end{aligned}$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.17 + 0.001 \cdot 2)} = 0.9885$$

$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17 + 0.001 \cdot 2}{(0.17 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)} \approx 0.5658$$

and also

$$\begin{aligned} m_{h(1)}(\theta_3) &= m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3) \\ &\approx 0.13 + (0.3489 \cdot 0.9851) + (0.0022 \cdot 0.4342) \\ &= 0.13 + 0.3437 + 0.0009 = 0.4746 \end{aligned}$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.13 + 0.001 \cdot 2}{(0.13 + 0.001 \cdot 2) + (0.00 + 0.001 \cdot 2)} = 0.9851$$

$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.13 + 0.001 \cdot 2}{(0.17 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)} \approx 0.4342$$

Hence, the final result of HPR approximation is

$$m_{h(1)}(\theta_1) = 0.0092, \quad m_{h(1)}(\theta_2) = 0.5161, \quad m_{h(1)}(\theta_3) = 0.4746$$

and we can easily verify that

$$m_{h(1)}(\theta_1) + m_{h(1)}(\theta_2) + m_{h(1)}(\theta_3) = 1.$$

The bba's obtained in each step are listed in Table III ($\epsilon = 0$) and Table IV ($\epsilon = 0.001$)

Table III
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 2 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k=3$	$k=2$	$k=1$
θ_1	0.0000	0.0000	0.0000
θ_2	0.1700	0.1700	0.5200
θ_3	0.1300	0.1300	0.4800
$\theta_1 \cup \theta_2$	0.2000	0.3500	0.0000
$\theta_1 \cup \theta_3$	0.2000	0.3500	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

Table IV
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 2 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k=3$	$k=2$	$k=1$
θ_1	0.0000	0.0000	0.0092
θ_2	0.1700	0.1700	0.5141
θ_3	0.1300	0.1300	0.4746
$\theta_1 \cup \theta_2$	0.2000	0.3489	0.0000
$\theta_1 \cup \theta_3$	0.2000	0.3489	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0022	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

When using $k-l-x$ approach, the results are in Table V.

Table V
EXPERIMENTAL RESULTS OF $k-l-x$ FOR EXAMPLE 2

Focal elements	$m(\cdot)$ obtained by $k-l-x$	
	$k=l=6$	$k=l=3$
θ_1	0.0000	0.0000
θ_2	0.1700	0.0000
θ_3	0.1300	0.0000
$\theta_1 \cup \theta_2$	0.2000	0.2857
$\theta_1 \cup \theta_3$	0.2000	0.2857
$\theta_2 \cup \theta_3$	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.4286

C. Example 3

Let's consider the following bba:

$$m(\theta_1) = 0, \quad m(\theta_2) = 0, \quad m(\theta_3) = 0.70,$$
$$m(\theta_1 \cup \theta_2) = 0, \quad m(\theta_1 \cup \theta_3) = 0,$$
$$m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$$

In this example, the mass assignments for all the focal elements with cardinality size 2 equal to zero. For HPR, when $\epsilon > 0$, $m(\theta_2 \cup \theta_3)$ will be divided equally and redistributed to $\{\theta_1 \cup \theta_2\}$, $\{\theta_1 \cup \theta_3\}$ and $\{\theta_2 \cup \theta_3\}$. Because the ratios are

$$R(\theta_1 \cup \theta_2, X(3)) = R(\theta_1 \cup \theta_3, X(3)) = R(\theta_2 \cup \theta_3, X(3)) = \frac{0.00 + 0.001 \cdot 3}{(0.00 + 0.001 \cdot 3) \cdot 3} = 0.3333$$

For HPR, when $\epsilon = 0$, it can not be executed directly. This can show the necessity for the using of ϵ.

The bba's obtained through $\text{HPR}_{\epsilon=0.001}$ at different steps are listed in Table VI

Table VI
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 3 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k=3$	$k=2$	$k=1$
θ_1	0.0000	0.0000	0.0503
θ_2	0.0000	0.0000	0.0503
θ_3	0.7000	0.7000	0.8994
$\theta_1 \cup \theta_2$	0.0000	0.1000	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.1000	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.1000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000

When using $k-l-x$ approach, the results are in Table VII.

Table VII
EXPERIMENTAL RESULTS OF $k-l-x$ FOR EXAMPLE 3

Focal elements	$m(\cdot)$ obtained by $k-l-x$	
	$k=l=6$	$k=l=3$
θ_1	0.0000	0.0000
θ_2	0.0000	0.0000
θ_3	0.7000	0.7000
$\theta_1 \cup \theta_2$	0.0000	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.3000

D. Example 4 (vacuous bba)

Let's consider the following bba:

$$m(\theta_1) = 0, \quad m(\theta_2) = 0, \quad m(\theta_3) = 0,$$
$$m(\theta_1 \cup \theta_2) = 0, \quad m(\theta_1 \cup \theta_3) = 0,$$
$$m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 1$$

In this example, the mass assignments for all the focal elements with cardinality size less than 3 equal to zero. For HPR, when $\epsilon > 0$, $m(\theta_1 \cup \theta_2 \cup \theta_3)$ will be divided equally and redistributed to $\{\theta_1 \cup \theta_2\}$, $\{\theta_1 \cup \theta_3\}$ and $\{\theta_2 \cup \theta_3\}$.

Similarly, the mass assignments for focal elements with cardinality of 2 obtained in intermediate step will be divided equally and redistributed to singletons. This is due to $\epsilon > 0$.

For HPR, when $\epsilon = 0$, it can not be executed directly. This can show the necessity for the using of ϵ. The bba's obtained through $\text{HPR}_{\epsilon=0.001}$ at different steps are listed in Table VIII.

Table VIII
EXPERIMENTAL RESULTS OF HPR FOR EXAMPLE 4 ($\epsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate baa		
	$k=3$	$k=2$	$k=1$
θ_1	0.0000	0.0000	0.3333
θ_2	0.0000	0.0000	0.3333
θ_3	0.0000	0.0000	0.3333
$\theta_1 \cup \theta_2$	0.0000	0.3333	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.3333	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.3333	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	1.0000	0.0000	0.0000

When using $k-l-x$ approach, the results are in Table IX.

Table IX
EXPERIMENTAL RESULTS OF $k-l-x$ FOR EXAMPLE 3

Focal elements	$m(\cdot)$ obtained by $k-l-x$	
	$k=l=6$	$k=l=3$
θ_1	0.0000	0.0000
θ_2	0.0000	0.0000
θ_3	0.0000	0.0000
$\theta_1 \cup \theta_2$	0.0000	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	1.0000	1.0000

From the results of Example 1 – Example 4, we can see that based on $k-l-x$, the users can control the number of focal elements but can not control the maximum cardinality of focal elements. Although based on $k-l-x$, the number of focal elements can be reduced, the focal elements with big cardinality might also be remained. This is not good for further reducing computational cost and not good for human to catch the meaning.

E. Example 5

More generally, an approximation method 1 (giving $m_1(.)$) is considered better than a method 2 (giving $m_2(.)$) if both conditions are fulfilled: 1) if Jousselme's distance of $m_1(.)$ to original bba $m(.)$ is smaller than the distance of $m_2(.)$ to original bba $m(.)$, i.e. $d(m_1, m) < d(m_2, m)$; 2) if the approximate non-specificity value $U(m_1)$ is closer (and lower) to the true non-specificity value $U(m)$ than $U(m_2)$, where Jousselme's distance is defined in [16], and non-specificity [17] is given by $U(m) = \sum_{A \subseteq \Theta} m(A) \log_2 |A|$.

In this example, we make a comparison between HPR (method 1) and k-additive approach (method 2). We consider the FoD $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ and we generate randomly $L = 30$ bba's by using the algorithm given below [15]:

Input: Θ : Frame of discernment;
N_{max}: Maximum number of focal elements
Output: Bel: Belief function (under the form of a bba, m)
Generate the power set of Θ: $\mathcal{P}(\Theta)$;

Generate a random permutation of $\mathcal{P}(\Theta) \rightarrow \mathcal{R}(\Theta)$;
Generate a integer between 1 and $N_{max} \rightarrow k$;
FOReach First k elements of $\mathcal{R}(\Theta)$ do
Generate a value within $[0,1] \rightarrow m'_k$;
END Normalize the vector $m'(.) = [m'_1, ..., m'_k] \rightarrow m(.)$ (that is $m(A_k) = m_k$);
Algorithm 1: Random generation of bba.

We compute and plot $d(m_1^j, m)$, $d(m_2^j, m)$, $U(m)$, $U(m_1^j)$ and $U(m_2^j)$ for several levels of approximation for $j = 1, 2, \ldots, L$ (where j is the index of the Monte-Carlo run). The results are shown in Fig. 3 and indicate clearly the superiority of HPR over the k-additive approach.

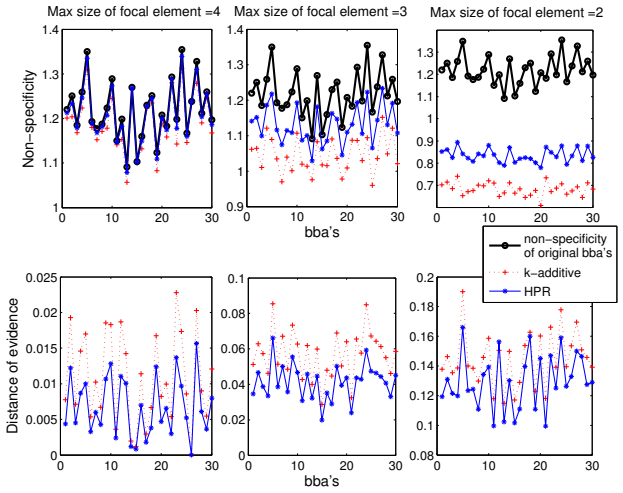


Figure 3. Illustration of Example 5.

We further use the Normalized Mean Square Error (NMSE) statistics defined by

$$NMSE_i = \frac{1}{L} \sum_{j=1}^{L} \frac{\left(U(m_i^j) - U(m)\right)^2}{\mathrm{Var}(\vec{e}_i)} \qquad (13)$$

to evaluate the global quality of the approximation of the non-specificity by HPR (if $i = 1$) and by k-additive method (if $i = 2$). $\vec{e}_i = [e_i^1, \ldots, e_i^j, \ldots, e_i^L]$ is the approximation error vector of method #i where $e_i^j = U(m_i^j) - U(m)$, for $j = 1, ..., L$. $\mathrm{Var}(\vec{e}_i)$ is the variance of $\vec{e}_i$. The NMSE results are given in Table X below.

Table X
NMSE RESULTS OF EXAMPLE 5

Max size of focal element	4	3	2
k-additive method	3.9003	21.8118	69.0191
HPR method	**3.9003**	**19.0264**	**61.9468**

Table X shows that HPR outperforms k-additive method since it provides a lower NMSE, which means that in terms of information loss, HPR is better (it generates less loss) than the k-additive approximation method.

VI. CONCLUSIONS

We have proposed a new interesting and useful hierarchical method, called HPR, to approximate any bba. The non-specificity degree can be easily controlled by the user. Some examples were provided to show how HPR works, and to show its rationality and advantage in comparison with some well-known bba approximation approaches. In future works, we will compare this HPR method with more bba approximation methods. In this paper, we have used only the distance of evidence and non-specificity as performance criteria. We plan to develop a more efficient evaluation criteria for capturing more aspects of the information expressed in a bba to measure the global performances of a method, and to design a better bba approximation approach (if possible).

REFERENCES

[1] G. Shafer, A Mathematical Theory of Evidence, Princeton, NJ: Princeton University, 1976.
[2] P. Smets, "Practical uses of belief functions", in K. B. Lskey and H. Prade, Editors, Uncertianty in Artificial Intelligence 15 (UAI 99), Stockholm, Sweden, pp. 612-621, 1999.
[3] B. Tessem, "Approximations for efficient computation in the theory of evidence", Artificial Intelligence, Vol. 61, no. 2, pp. 315-329, June 1993.
[4] F. Voorbraak, "A computationally efficient approximation of Dempster-Shafer theory", Int. J. Man-Machine Studies, Vol. 30, pp. 525-536, 1989.
[5] F. Smarandache, J. Dezert (Editors), Applications and Advances of DSmT for Information Fusion (Vol 3), Rehoboth, NM: American Research Press, 2009. http://www.gallup.unm.edu/˜smarandache/DSmT-book3.pdf.
[6] R. Kennes, "Computational aspects of the Möbius transform of graphs", IEEE Transactions on SMC, Vol. 22, pp. 201-223, 1992.
[7] J.A. Barnett, "Computational methods for a mathematical theory of evidence", in Proceedings of IJCAI-81, Vancouver, pp. 868-875, 1981.
[8] G. Shafer, R. Logan, "Implementing Dempster's rule for hierarchical evidence", Artificial Intelligence, Vol. 33, pp. 271-298, 1987.
[9] S. Moral, A. Salmeron, "A Monte Carlo algorithm for combining Dempster-Shafer belief based on approximate pre-computation", in A. Hunter and S. Pearsons, Editors, Symbolic and quantitative approaches to reasoning and uncertainty (ECSQARU'99), London, UK, pp. 305-315, 1999.
[10] D. Dubois, H. Prade, "An alternative approach to the handling of subnormal possiblity distributions", Fuzzy Sets and Systems, Vol. 24, pp. 123-126, 1987.
[11] M. Grabisch, "Upper approximation of non-additive measures by k-additive measures - the case of belief functions", in Proc. of 1st Int. Symp. on Imprecise Proba. and their applications, Ghent, Belgium, June 1999.
[12] T. Burger, F. Cuzzolin, "Two k-additive generalizations of the pignistic transform", submitted in 2011 to Fuzzy Sets and Systems, available on line: http://cms.brookes.ac.uk/staff/FabioCuzzolin/files/fss11kadditive.pdf
[13] T. Denœux, "Inner and outer approximation of belief structures using a hierarchical clustering approach", Int. J. of Uncertainty, Fuzziness, and Knowledge-based Systems, Vol. 9, no. 4, pp. 437-460, 2001.
[14] T. Burger, "Defining new approximations of belief functions by means of Dempster's combination", in Proc. of the 1st International Workshop on the Theories of Belief Functions (WTBF 2010), Brest, France, March 31st - April 2nd, 2010.
[15] A.-L. Jousselme, P. Maupin, "On some properties of distances in evidence theory", in Proc. of the 1st Workshop on Theory of Belief Functions(WTBF2010), Brest, France, March 31st - April 2nd, 2010.
[16] A.-L. Jousselme, D. Grenier, E. Bosse, "A new distance between two bodies of evidence", Information Fusion, Vol. 2, no. 2, pp. 91-101, 2001.
[17] D. Dubois, H. Prade, "A note on measures of specificity for fuzzy sets", International Journal of General Systems, Vol. 10, no. 4, pp. 279-283, 1985.

Hierarchical DSmP Transformation for Decision-Making under Uncertainty

Jean Dezert
Deqiang Han
Zhun-ga Liu
Jean-Marc Tacnet

Originally published as Dezert J., Han D., Liu Z., Tacnet J.-M., *Hierarchical DSmP transformation for decision-making under uncertainty,* in Proc. of Fusion 2012, Singapore, July 2012, and reprinted with permission.

***Abstract*—Dempster-Shafer evidence theory is widely used for approximate reasoning under uncertainty; however, the decision-making is more intuitive and easy to justify when made in the probabilistic context. Thus the transformation to approximate a belief function into a probability measure is crucial and important for decision-making based on evidence theory framework. In this paper we present a new transformation of any general basic belief assignment (bba) into a Bayesian belief assignment (or subjective probability measure) based on new proportional and hierarchical principle of uncertainty reduction. Some examples are provided to show the rationality and efficiency of our proposed probability transformation approach.**

Keywords: Belief functions, probabilistic transformation, DSmP, uncertainty, decision-making.

I. Introduction

Dempster-Shafer evidence theory (DST) [1] proposes a mathematical framework for approximate reasoning under uncertainty thanks to belief functions. Thus it is widely used in many fields of information fusion. As any theory, DST is not exempt of drawbacks and limitations, like its inconsistency with the probability calculus, its complexity and the miss of a clear decision-making process. Aside these weaknesses, the use of belief functions remains flexible and appealing for modeling and dealing with uncertain and imprecise information. That is why several modified models and rules of combination of belief functions were proposed to resolve some of the drawbacks of the original DST. Among the advances in belief function theories, one can underline the transferable belief model (TBM) [2] proposed by Smets, and more recently the DSmT [3] proposed by Dezert and Smarandache.

The ultimate goal of approximate reasoning under uncertainty is usually the decision-making. Although the decision-making can be done based on evidence expressed by a belief function [4], the decision-making is better established in a probabilistic context: decisions can be evaluated by assessing their ability to provide a winning strategy on the long run in a game theory context, or by maximizing return in a utility theory framework. Thus to take a decision, it is usually preferred to transform (approximate) a belief function into a probability measure. So the quality of such probability transformation is crucial for the decision-making in the evidence theory. The research on probability transformation has attracted more attention in recent years.

The classical probability transformation in evidence theory is the pignistic probability transformation (PPT) [2] in TBM. TBM has two levels: the credal level, and the pignistic level. Beliefs are entertained, combined and updated at the credal level while the decision making is done at the pignistic level. PPT maps the beliefs defined on subsets to the probability defined on singletons. In PPT, belief assignments for a compound focal element are equally assigned to the singletons included. In fact, PPT is designed according to the principle of minimal commitment, which is somehow related with uncertainty maximization.

Other researchers also proposed some modified probability transformation approaches [5]–[13] to assign the belief assignments of compound focal elements to the singletons according to some ratio constructed based on some available information. The representative transformations include Sudano's probability transformations [8] and Cuzzolin's intersection probability transformation [13], etc. In the framework of DSmT, another probability transformation approach was proposed, which is called DSmP [9]. DSmP takes into account both the values of the masses and the cardinality of focal elements in the proportional redistribution process. DSmP can also be used in both DSmT and DST. For a probability transformation, it is always evaluated by using probabilistic information content (PIC) [5] (PIC being the dual form of Shannon entropy), although it is not enough or comprehensive [14]. A probability transformation providing a high probabilistic information content (PIC) is preferred in fact for decision-making since naturally it is always easier to take a decision when the uncertainty is reduced.

In this paper we propose a new probability transformation, which can output a probability with high but not exaggerated PIC. The new approach, called HDSmP (standing for Hierarchical DSmP) is implemented hierarchically and it fully utilize the information provided by a given belief function. Succinctly, for a frame of discernment (FOD) with size n, for $k = n$ down to $k = 2$, the following step is repeated: the belief assignment of a focal element with size k is proportionally redistributed to the focal elements with size $k - 1$. The proportion is defined by the ratio among mass assignments of focal elements with size $k-1$. A parameter ϵ is introduced in the formulas to avoid division by zero and warranty numerical

robustness of the result. HDSmP corresponds to the last step of the hierarchical proportional redistribution method for basic belief assignment (bba) approximation presented briefly in [16] and in more details in [17]. Some examples are given at the end of this paper to illustrate our proposed new probability transformation approach. Comparisons of our new HDSmP approach with the other well-known approaches with related analyses are also provided.

II. Evidence theory and probability transformations

A. Brief introduction of evidence theory

In Dempster-Shafer theory [1], the elements in the frame of discernment (FOD) Θ are mutually exclusive. Suppose that 2^Θ represents the powerset of FOD, and one defines the function $m : 2^\Theta \rightarrow [0,1]$ as the basic belief assignment (bba), also called mass function satisfying:

$$\sum\nolimits_{A\subseteq\Theta} m(A) = 1,\ m(\emptyset) = 0 \tag{1}$$

Belief function (Bel) and plausibility function (Pl) are defined below, respectively:

$$Bel(A) = \sum\nolimits_{B\subseteq A} m(B) \tag{2}$$

$$Pl(A) = \sum\nolimits_{A\cap B\neq\emptyset} m(B) \tag{3}$$

Suppose that $m_1, m_2, ..., m_n$ are n mass functions, Dempster's rule of combination is defined in (4):

$$m(A) = \begin{cases} 0, & A = \emptyset \\ \dfrac{\sum\limits_{\cap A_i = A} \prod\limits_{1\leq i\leq n} m_i(A_i)}{\sum\limits_{\cap A_i \neq \emptyset} \prod\limits_{1\leq i\leq n} m_i(A_i)}, & A \neq \emptyset \end{cases} \tag{4}$$

Dempster's rule of combination is used in DST to accomplish the fusion of bodies of evidence (BOEs). However, the final goal for decision-level information fusion is decision making. The beliefs should be transformed into probabilities, before the probability-based decision-making. Although there are also some research works on making decision directly based on belief function or bba [4], probability-based decision methods are more intuitive and have become the current trend to decide under uncertainty from approximate reasoning theories [15]. Some existing and well-known probability transformation approaches are briefly reviewed in the next section.

B. Probability transformations used in DST framework

A probability transformation (or briefly a "probabilization") is a mapping $PT : Bel_\Theta \rightarrow P_\Theta$, where Bel_Θ means the belief function defined on Θ and P_Θ represents a probability measure (in fact a probability mass function, pmf) defined on Θ. Thus the probability transformation assigns a Bayesian belief function (i.e. probability measure) to any general (i.e. non-Bayesian) belief function. It is a reason why the transformations from belief functions to probability distributions are sometimes called also Bayesian transformations.

The major probability transformation approaches used so far are:

a) Pignistic transformation

The classical pignistic probability was proposed by Smets [2] in his TBM framework which is a subjective and a non-probabilistic interpretation of DST. It extends the evidence theory to the open-world propositions and it has a range of tools including discounting and conditioning to handle belief functions. At the credal level of TBM, beliefs are entertained, combined and updated. While at the pignistic level, beliefs are used to make decisions by resorting to pignistic probability transformation (PPT). The pignistic probability obtained is always called betting commitment probability (in short, BetP). The basic idea of pignistic transformation consists of transferring the positive belief of each compound (or nonspecific) element onto the singletons involved in that compound element split by the cardinality of the proposition when working with normalized bba's.

Suppose that $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$ is the FOD. The PPT for the singletons is defined as [2]:

$$\text{BetP}_m(\theta_i) = \sum_{\theta_i\in B,\ B\in 2^\Theta} \frac{m(B)}{|B|} \tag{5}$$

PPT is designed according to an idea similar to uncertainty maximization. In PPT, masses are not assigned discriminately to different singletons involved. For information fusion, the aim is to reduce the degree of uncertainty and to gain a more consolidated and reliable decision result. High uncertainty in PPT might not be helpful for the decision. To overcome this, some typical modified probability transformation approaches were proposed which are summarized below.

b) Sudano's probabilities

Sudano [8] proposed Probability transformation proportional to Plausibilities (PrPl), Probability transformation proportional to Beliefs (PrBel), Probability transformation proportional to the normalized Plausibilities (PrNPl), Probability transformation proportional to all Plausibilities (PraPl) and Hybrid Probability transformation (PrHyb), respectively. As suggested by their names, different kinds of mappings were used. For the belief function defined on the FOD $\Theta = \{\theta_1, ..., \theta_n\}$, they are respectively defined by

$$\text{PrPl}(\theta_i) = Pl(\{\theta_i\}) \cdot \sum_{Y\in 2^\Theta, \theta_i\in Y} \frac{m(Y)}{\sum\limits_{\cup_j \theta_j = Y} Pl(\{\theta_j\})} \tag{6}$$

$$\text{PrBel}(\theta_i) = Bel(\{\theta_i\}) \cdot \sum_{Y\in 2^\Theta, \theta_i\in Y} \frac{m(Y)}{\sum\limits_{\cup_j \theta_j = Y} Bel(\{\theta_j\})} \tag{7}$$

$$\text{PrNPl}(\theta_i) = \frac{Pl(\{\theta_i\})}{\sum\limits_j Pl(\{\theta_j\})} \tag{8}$$

$$\text{PraPl}(\theta_i) = Bel(\{\theta_i\}) + \frac{1 - \sum_j Bel(\{\theta_j\})}{\sum_j Pl(\{\theta_j\})} \cdot Pl(\{\theta_i\}) \tag{9}$$

$$\text{PrHyb}(\theta_i) = \text{PraPl}(\theta_i) \cdot \sum_{Y \in 2^\Theta, \theta_i \in Y} \frac{m(Y)}{\sum\limits_{\cup_j \theta_j = Y} \text{PraPl}(\theta_j)} \quad (10)$$

c) Cuzzolin's intersection probability

From a geometric interpretation of Dempster's rule of combination, an intersection probability measure was proposed by Cuzzolin [12] from the proportional repartition of the total nonspecific mass (TNSM) for each contribution of the nonspecific masses involved.

$$\text{CuzzP}(\theta_i) = m(\{\theta_i\}) + \frac{Pl(\{\theta_i\}) - m(\{\theta_i\})}{\sum_j (Pl(\{\theta_j\}) - m(\{\theta_j\}))} \cdot \text{TNSM} \quad (11)$$

where

$$\text{TNSM} = 1 - \sum\nolimits_j m(\{\theta_j\}) = \sum\nolimits_{A \in 2^\Theta, |A|>1} m(A) \quad (12)$$

d) DSmP transformation

DSmP proposed recently by Dezert and Smarandache is defined as follows:

$$\text{DSmP}_\epsilon(\theta_i) = m(\{\theta_i\}) + (m(\{\theta_i\}) + \epsilon) \cdot \sum_{\substack{X \in 2^\Theta \\ \theta_i \subset X \\ |X| \geq 2}} \frac{m(X)}{\sum_{\substack{Y \in 2^\Theta \\ Y \subset X \\ |Y|=1}} m(Y) + \epsilon \cdot |X|} \quad (13)$$

In DSmP, both the mass assignments and the cardinality of focal elements are used in the proportional redistribution process. The parameter of ϵ is used to adjust the effect of focal element's cardinality in the proportional redistribution, and to make DSmP defined and computable when encountering zero masses. DSmP made an improvement compared with Sudano's, Cuzzolin's and PPT formulas, in that DSmP mathematically makes a more judicious redistribution of the ignorance masses to the singletons involved and thus increases the PIC level of the resulting approximation. Moreover, DSmP works for both theories of DST and DSmT.

There are still some other definitions on modified PPT such as the iterative and self-consistent approach PrScP proposed by Sudano in [5], and a modified PrScP in [11]. Although the aforementioned probability transformation approaches are different, they are all evaluated according to the degree of uncertainty. The classical evaluation criteria for a probability transformation are the following ones:

1) Normalized Shannon Entropy

Suppose that $P(\theta)$ is a probability mass function (pmf), where $\theta \in \Theta$, $|\Theta| = N$ and the $|\Theta|$ represents the cardinality of the FOD Θ. An evaluation criterion for the pmf obtained from different probability transformation is as follows [12]:

$$\text{E}_\text{H} = \frac{-\sum\limits_{\theta \in \Theta} P(\theta) \log_2(P(\theta))}{\log_2 N} \quad (14)$$

i.e., the ratio of Shannon entropy and the maximum of Shannon entropy for $\{P(\theta)|\theta \in \Theta\}$,$|\Theta| = N$. Clearly E_H is normalized. The larger E_H is, the larger the degree of uncertainty is. The smaller E_H is, the smaller the degree of uncertainty is. When E_H= 0, one hypothesis will have probability 1 and the rest with zero probabilities. Therefore the agent or system can make decision without error. When E_H= 1, it is impossible to make a correct decision, because $P(\theta)$, for all $\theta \in \Theta$ are equal.

2) Probabilistic Information Content

Probabilistic Information Content (PIC) criterion [5] is an essential measure in any threshold-driven automated decision system. The PIC value of a pmf obtained from a probability transformation indicates the level of the total knowledge one has to draw a correct decision.

$$\text{PIC}(P) = 1 + \frac{1}{\log_2 N} \cdot \sum_{\theta \in \Theta} P(\theta) \log_2(P(\theta)) \quad (15)$$

Obviously, $\text{PIC} = 1 - \text{E}_\text{H}$. The PIC is the dual of the normalized Shannon entropy. A PIC value of zero indicates that the knowledge to take a correct decision does not exist (all hypotheses have equal probabilities, i.e., one has the maximal entropy).

Less uncertainty means that the corresponding probability transformation result is better to help to take a decision. According to such a simple and basic idea, the probability transformation approach should attempt to enlarge the belief differences among all the propositions and thus to achieve a more reliable decision result.

III. The Hierarchical DSmP transformation

In this paper, we propose a novel probability transformation approach called hierarchical DSmP (HDSmP), which provides a new way to reduce step by step the mass committed to uncertainties until to obtain an approximate measure of subjective probability, i.e. a so-called Bayesian bba in [1]. It must be noticed that this procedure can be stopped at any step in the process and thus it allows to reduce the number of focal elements in a given bba in a consistent manner to diminish the size of the core of a bba and thus reduce the complexity (if needed) when applying also some complex rules of combinations. We present here the general principle of hierarchical and proportional reduction of uncertainties in order to finally obtain a Bayesian bba. The principle of redistribution of uncertainty to more specific elements of the core at any given step of the process follows the proportional redistribution already proposed in the (non hierarchical) DSmP transformation proposed recently in [3].

Let's first introduce two new notations for convenience and for concision:

1) Any element of cardinality $1 \leq k \leq n$ of the power set 2^Θ will be denoted, by convention, by the generic notation $X(k)$. For example, if $\Theta = \{A, B, C\}$, then $X(2)$ denotes the following partial uncertainties $A \cup B$, $A \cup C$ or $B \cup C$, and $X(3)$ denotes the total uncertainty $A \cup B \cup C$.
2) The proportional redistribution factor (ratio) of width s involving elements Y and X of the powerset is defined as (for $X \neq \emptyset$ and $Y \neq \emptyset$)

$$R_s(Y,X) \triangleq \frac{m(Y)+\epsilon\cdot|X|}{\sum_{\substack{Y\subset X\\ |X|-|Y|=s}}(m(Y)+\epsilon\cdot|X|)} \qquad (16)$$

where ϵ is a small positive number introduced here to deal with particular cases where $\sum_{\substack{Y\subset X\\ |X|-|Y|=s}} m(Y)=0$. In HDSmP, we just need to use the proportional redistribution factors of width $n=1$, and so we will just denote $R(Y,X)\triangleq R_1(Y,X)$ by convention.

The HDSmP transformation is obtained by a step by step (recursive) proportional redistribution of the mass $m(X(k))$ of a given uncertainty $X(k)$ (partial or total) of cardinality $2\leq k\leq n$ to all the least specific elements of cardinality $k-1$, i.e. to all possible $X(k-1)$, until $k=2$ is reached. The proportional redistribution is done from the masses of belief committed to $X(k-1)$ as done classically in DSmP transformation. Mathematically, HDSmP is defined for any $X(1)\in\Theta$, i.e. any $\theta_i\in\Theta$ by

$$HDSmP(X(1)) = m(X(1)) + \sum_{\substack{X(2)\supset X(1)\\ X(1),X(2)\in 2^\Theta}} [m_h(X(2))\cdot R(X(1),X(2))] \qquad (17)$$

where the "hierarchical" masses $m_h(.)$ are recursively (backward) computed as follows:

$$m_h(X(n-1)) = m(X(n-1)) + \sum_{\substack{X(n)\supset X(n-1)\\ X(n),X(n-1)\in 2^\Theta}} [m(X(n))\cdot R(X(n-1),X(n))]$$

$$m_h(A)=m(A),\forall|A|<n-1 \qquad (18)$$

$$m_h(X(n-2)) = m(X(n-2)) + \sum_{\substack{X(n-1)\supset X(n-2)\\ X(n-2),X(n-1)\in 2^\Theta}} [m_h(X(n-1))\cdot R(X(n-2),X(n-1))]$$

$$m_h(A)=m(A),\forall|A|<n-2$$
$$\vdots \qquad (19)$$

$$m_h(X(2)) = m(X(2)) + \sum_{\substack{X(3)\supset X(2)\\ X(3),X(2)\in 2^\Theta}} [m_h(X(3))\cdot R(X(2),X(3))]$$

$$m_h(A)=m(A),\forall|A|<2 \qquad (20)$$

Actually, it is worth to note that $X(n)$ is in fact unique and it corresponds only to the full ignorance $\theta_1\cup\theta_2\cup\ldots\cup\theta_n$. Therefore, the expression of $m_h(X(n-1))$ in Eq. (18) just simplifies as

$$m_h(X(n-1)) = m(X(n-1)) + m(X(n))\cdot R(X(n-1),X(n))$$

Because of the full proportional redistribution of the masses of uncertainties to the elements least specific involved in these uncertainties, no mass of belief is lost during the step by step hierarchical process and thus one gets finally a Bayesian bba satisfying $\sum_{X(1)\in 2^\Theta} HDSmP(X(1))=1$.

IV. Examples

In this section we show in details how HDSmP can be applied on very simple different examples. So let's examine the three following examples based on a simple 3D frame of discernment $\Theta=\{\theta_1,\theta_2,\theta_3\}$ satisfying Shafer's model.

A. Example 1

Let's consider the following bba:

$$m(\theta_1)=0.10,\quad m(\theta_2)=0.17,\quad m(\theta_3)=0.03,$$
$$m(\theta_1\cup\theta_2)=0.15,\quad m(\theta_1\cup\theta_3)=0.20,$$
$$m(\theta_2\cup\theta_3)=0.05,\quad m(\theta_1\cup\theta_2\cup\theta_3)=0.30.$$

We apply HDSmP with $\epsilon=0$ in this example because there is no mass of belief equal to zero. It can be verified that the result obtained with a small positive ϵ parameter remains (as expected) numerically very close to the result obtained with $\epsilon=0$. This verification is left to the reader.

The first step of HDSmP consists in redistributing back $m(\theta_1\cup\theta_2\cup\theta_3)=0.30$ committed to the full ignorance to the elements $\theta_1\cup\theta_2$, $\theta_1\cup\theta_3$ and $\theta_2\cup\theta_3$ only, because these elements are the only elements of cardinality 2 that are included in $\theta_1\cup\theta_2\cup\theta_3$. Applying the Eq. (18) with $n=3$, one gets when $X(2)=\theta_1\cup\theta_2$, $\theta_1\cup\theta_3$ and $\theta_1\cup\theta_2$ the following masses

$$m_h(\theta_1\cup\theta_2) = m(\theta_1\cup\theta_2)+m(X(3))\cdot R(\theta_1\cup\theta_2,X(3)) = 0.15+(0.3\cdot 0.375)=0.2625$$

because $R(\theta_1\cup\theta_2,X(3))=\frac{0.15}{0.15+0.20+0.05}=0.375$. Similarly, one gets

$$m_h(\theta_1\cup\theta_3) = m(\theta_1\cup\theta_3)+m(X(3))\cdot R(\theta_1\cup\theta_3,X(3)) = 0.20+(0.3\cdot 0.5)=0.35$$

because $R(\theta_1\cup\theta_3,X(3))=\frac{0.20}{0.15+0.20+0.05}=0.5$, and also

$$m_h(\theta_2\cup\theta_3) = m(\theta_2\cup\theta_3)+m(X(3))\cdot R(\theta_2\cup\theta_3,X(3)) = 0.05+(0.3\cdot 0.125)=0.0875$$

because $R(\theta_2\cup\theta_3,X(3))=\frac{0.05}{0.15+0.20+0.05}=0.125$.

Now, we go to the next step of HDSmP and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1\cup\theta_2$, $\theta_1\cup\theta_3$ and $\theta_2\cup\theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use directly HDSmP in Eq. (17) for doing this as follows:

$$\begin{aligned} HDSmP(\theta_1) &= m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3) \\ &\approx 0.10 + (0.2625 \cdot 0.3703) + (0.35 \cdot 0.7692) \\ &= 0.10 + 0.0972 + 0.2692 = 0.4664 \end{aligned}$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.10}{0.10 + 0.17} \approx 0.3703$$

$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.10}{0.10 + 0.03} \approx 0.7692$$

Similarly, one gets

$$\begin{aligned} HDSmP(\theta_2) &= m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3) \\ &\approx 0.10 + (0.2625 \cdot 0.6297) + (0.0875 \cdot 0.85) \\ &= 0.17 + 0.1653 + 0.0744 = 0.4097 \end{aligned}$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17}{0.10 + 0.17} \approx 0.6297$$

$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17}{0.17 + 0.03} = 0.85$$

and also

$$\begin{aligned} HDSmP(\theta_3) &= m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3) \\ &\approx 0.03 + (0.35 \cdot 0.2307) + (0.0875 \cdot 0.15) \\ &= 0.03 + 0.0808 + 0.0131 = 0.1239 \end{aligned}$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.03}{0.10 + 0.03} \approx 0.2307$$

$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.03}{0.17 + 0.03} = 0.15$$

Hence, the final result of HDSmP transformation is:

$$\begin{aligned} &HDSmP(\theta_1) = 0.4664, \quad HDSmP(\theta_2) = 0.4097, \\ &HDSmP(\theta_3) = 0.1239. \end{aligned}$$

and we can easily verify that

$$HDSmP(\theta_1) + HDSmP(\theta_2) + HDSmP(\theta_3) = 1.$$

The procedure can be illustrated in Fig. 1 below.

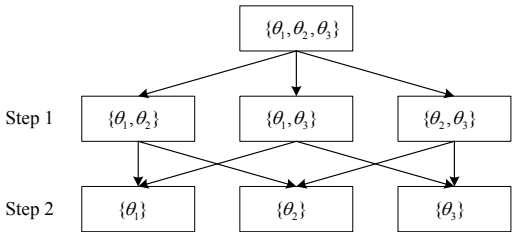


Figure 1. Illustration of Example 1

Table I
EXPERIMENTAL RESULTS FOR EXAMPLE 1.

Approaches	Propositions			E_H
	θ_1	θ_2	θ_3	
BetP	0.3750	0.3700	0.2550	0.9868
PrPl	0.4045	0.3681	0.2274	0.9747
PrBel	0.4094	0.4769	0.1137	0.8792
DSmP_0	0.4094	0.4769	0.1137	0.8792
DSmP_0.001	0.4094	0.4769	0.1137	0.8792
HDSmP_0	0.4664	0.4097	0.1239	0.8921
HDSmP_0.001	0.4664	0.4097	0.1239	0.8921

The classical DSmP transformation [3] and the other transformations (BetP [2], PrBel and PrPl [8]) are compared with HDSmP for this example in Table I. It can be seen in Table I that the normalized entropy E_H of HDSmP is relatively small but not too small among all the probability transformations used. In fact it is normal that the entropy drawn form HDSmP is a bit bigger than the entropy drawn from DSmP, because there is a "dilution" of uncertainty in the step-by-step redistribution, whereas such dilution of uncertainty is absent in the direct DSmP transformation.

B. Example 2

Let's consider the following bba:

$$\begin{aligned} &m(\theta_1) = 0, \quad m(\theta_2) = 0.17, \quad m(\theta_3) = 0.13, \\ &m(\theta_1 \cup \theta_2) = 0.20, \quad m(\theta_1 \cup \theta_3) = 0.20, \\ &m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30 \end{aligned}$$

The first step of HDSmP consists in redistributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, and $\theta_1 \cup \theta_3$ only, because these elements are the only elements of cardinality 2 that are included in $\theta_1 \cup \theta_2 \cup \theta_3$. Applying the Eq. (18) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses

$$\begin{aligned} m_h(\theta_1 \cup \theta_2) &= m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3)) \\ &= 0.20 + (0.3 \cdot 0.5) = 0.35 \end{aligned}$$

because $R(\theta_1 \cup \theta_2, X(3)) = \frac{0.20}{0.20+0.20+0.00} = 0.5$.
Similarly, one gets

$$\begin{aligned} m_h(\theta_1 \cup \theta_3) &= m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3)) \\ &= 0.20 + (0.3 \cdot 0.5) = 0.35 \end{aligned}$$

because $R(\theta_1 \cup \theta_3, X(3)) = \frac{0.20}{0.20+0.20+0.00} = 0.5$, and also

$$\begin{aligned} m_h(\theta_2 \cup \theta_3) &= m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3)) \\ &= 0.00 + (0.3 \cdot 0.0) = 0 \end{aligned}$$

because $R(\theta_2 \cup \theta_3, X(3)) = \frac{0.0}{0.20+0.20+0.00} = 0$.

Now, we go to the next and last step of HDSmP principle, and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3.

We use directly HDSmP in Eq. (17) for doing this as follows:

$$\begin{aligned} HDSmP(\theta_1) &= m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3) \\ &\approx 0.00 + (0.35 \cdot 0.00) + (0.35 \cdot 0.00) \\ &= 0.00 + 0.00 + 0.00 = 0 \end{aligned}$$

because

$$R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.00}{0.00 + 0.17} = 0.00$$

$$R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.00}{0.00 + 0.13} = 0.00$$

Similarly, one gets

$$\begin{aligned} HDSmP(\theta_2) &= m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3) \\ &\approx 0.17 + (0.35 \cdot 1) + (0.00 \cdot 0.5667) \\ &= 0.17 + 0.35 + 0.00 = 0.52 \end{aligned}$$

because

$$R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17}{0.00 + 0.17} = 1$$

$$R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17}{0.17 + 0.13} \approx 0.5667$$

and also

$$\begin{aligned} HDSmP(\theta_3) &= m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3) \\ &\approx 0.13 + (0.35 \cdot 1) + (0.00 \cdot 0.4333) \\ &= 0.13 + 0.35 + 0.00 = 0.48 \end{aligned}$$

because

$$R(\theta_3, \theta_1 \cup \theta_3) = \frac{0.13}{0.13 + 0.00} = 1$$

$$R(\theta_3, \theta_2 \cup \theta_3) = \frac{0.13}{0.17 + 0.13} \approx 0.4333$$

Hence, the final result of HDSmP transformation is:

$$HDSmP(\theta_1) = 0.4664, \qquad HDSmP(\theta_2) = 0.4097,$$
$$HDSmP(\theta_3) = 0.1239.$$

and we can easily verify that

$$HDSmP(\theta_1) + HDSmP(\theta_2) + HDSmP(\theta_3) = 1.$$

The HDSmP procedure of Example 2 with $\epsilon = 0$ is Fig. 2. The HDSmP procedure of Example 2 with $\epsilon > 0$ is the same as that illustrated in Fig. 1. When one takes $\epsilon > 0$, there exist masses redistributed to $\{\theta_2 \cup \theta_3\}$. If one takes $\epsilon = 0$, there is no mass edistributed to $\{\theta_2 \cup \theta_3\}$. That's the difference between Fig. 1 and Fig. 2.

Let's suppose that one takes $\epsilon = 0.001$, then the HDSmP calculation procedure is as follows:

• Step 1: The first step of HDSmP consists in distributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$. Applying the formula

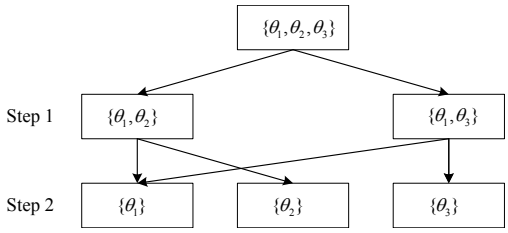


Figure 2. Illustration of Example 2.

(III) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses

$$\begin{aligned} m_h(\theta_1 \cup \theta_2) &= m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3)) \\ &= 0.20 + (0.3 \cdot 0.4963) = 0.3489 \end{aligned}$$

because

$$\begin{aligned} R(\theta_1 \cup \theta_2, X(3)) &= \frac{0.20 + 0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)} \\ &= 0.4963 \end{aligned}$$

$$\begin{aligned} m_h(\theta_1 \cup \theta_3) &= m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3)) \\ &= 0.20 + (0.3 \cdot 0.4963) = 0.3489 \end{aligned}$$

because

$$\begin{aligned} R(\theta_1 \cup \theta_2, X(3)) &= \frac{0.20 + 0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)} \\ &= 0.4963 \end{aligned}$$

$$\begin{aligned} m_h(\theta_2 \cup \theta_3) &= m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3)) \\ &= 0.00 + (0.3 \cdot 0.0073) = 0.0022 \end{aligned}$$

because

$$\begin{aligned} R(\theta_2 \cup \theta_3, X(3)) &= \frac{0.001 \cdot 3}{(0.20 + 0.001 \cdot 3) \cdot 2 + (0.00 + 0.001 \cdot 3)} \\ &= 0.0073 \end{aligned}$$

• Next step: one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use directly HDSmP in Eq. (17) for doing this as follows:

$$\begin{aligned} HDSmP(\theta_1) &= m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3) \\ &\approx 0.00 + (0.3489 \cdot 0.0115) + (0.3489 \cdot 0.0149) \\ &= 0.00 + 0.0040 + 0.0052 = 0.0092 \end{aligned}$$

because

$$\begin{aligned} R(\theta_1, \theta_1 \cup \theta_2) &= \frac{0.00 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.17 + 0.001 \cdot 2)} \\ &= 0.0115 \end{aligned}$$

$$\begin{aligned} R(\theta_1, \theta_1 \cup \theta_3) &= \frac{0.00 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)} \\ &= 0.0149 \end{aligned}$$

Similarly, one gets

$$\begin{aligned} HDSmP(\theta_2) &= m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3) \\ &\approx 0.17 + (0.3489 \cdot 0.9885) + (0.0022 \cdot 0.5658) \\ &= 0.17 + 0.3449 + 0.0012 = 0.5161 \end{aligned}$$

because

$$\begin{aligned} R(\theta_2, \theta_1 \cup \theta_2) &= \frac{0.17 + 0.001 \cdot 2}{(0.00 + 0.001 \cdot 2) + (0.17 + 0.001 \cdot 2)} \\ &= 0.9885 \end{aligned}$$

$$\begin{aligned} R(\theta_2, \theta_2 \cup \theta_3) &= \frac{0.17 + 0.001 \cdot 2}{(0.17 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)} \\ &\approx 0.5658 \end{aligned}$$

and also

$$\begin{aligned} HDSmP(\theta_3) &= m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3) \\ &\quad + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3) \\ &\approx 0.13 + (0.3489 \cdot 0.9851) + (0.0022 \cdot 0.4342) \\ &= 0.13 + 0.3437 + 0.0009 = 0.4746 \end{aligned}$$

because

$$\begin{aligned} R(\theta_3, \theta_1 \cup \theta_3) &= \frac{0.13 + 0.001 \cdot 2}{(0.13 + 0.001 \cdot 2) + (0.00 + 0.001 \cdot 2)} \\ &= 0.9851 \end{aligned}$$

$$\begin{aligned} R(\theta_3, \theta_2 \cup \theta_3) &= \frac{0.13 + 0.001 \cdot 2}{(0.17 + 0.001 \cdot 2) + (0.13 + 0.001 \cdot 2)} \\ &\approx 0.4342 \end{aligned}$$

Hence, the final result of HDSmP transformation is:

$$\begin{aligned} &HDSmP(\theta_1) = 0.0092, \quad HDSmP(\theta_2) = 0.5161, \\ &HDSmP(\theta_3) = 0.4746. \end{aligned}$$

and we can easily verify that

$$HDSmP(\theta_1) + HDSmP(\theta_2) + HDSmP(\theta_3) = 1.$$

We also calculate some other probability transformations and the results are listed in Table II.

Table II
EXPERIMENTAL RESULTS FOR EXAMPLE 2.

Approaches	Propositions			E_H
	θ_1	θ_2	θ_3	
BetP	0.3000	0.3700	0.3300	0.9966
PrPl	0.3125	0.3683	0.3192	0.9975
PrBel	NaN	NaN	NaN	NaN
DSmP_0	0.0000	0.5400	0.4600	0.6280
DSmP_0.001	0.0037	0.5381	0.4582	0.6479
HDSmP_0	0.0000	0.5200	0.4800	0.6302
HDSmP_0.001	0.0092	0.5161	0.4746	0.6720

It can be seen in Table II that the normalized entropy E_H of HDSmP is relatively small but not too small among all the probability transformations used.

C. Example 3

Let's consider the following bba:

$$\begin{aligned} &m(\theta_1) = 0, \quad m(\theta_2) = 0, \quad m(\theta_3) = 0.70, \\ &m(\theta_1 \cup \theta_2) = 0, \quad m(\theta_1 \cup \theta_3) = 0, \\ &m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30 \end{aligned}$$

In this example, the mass assignments for all the focal elements with cardinality size 2 equal to zero. For HDSmP, when $\epsilon > 0$, $m(\theta_2 \cup \theta_3)$ will be divided equally and redistributed to $\{\theta_1 \cup \theta_2\}$, $\{\theta_1 \cup \theta_3\}$ and $\{\theta_2 \cup \theta_3\}$. Because the ratios are

$$\begin{aligned} R(\theta_1 \cup \theta_2, X(3)) &= R(\theta_1 \cup \theta_3, X(3)) = R(\theta_2 \cup \theta_3, X(3)) \\ &= \frac{0.00 + 0.001 \cdot 3}{(0.00 + 0.001 \cdot 3) \cdot 3} = 0.3333 \end{aligned}$$

One sees that with the parameter $\epsilon = 0$, HDSmP cannot be computed (division by zero) and that is why it is necessary to use $\epsilon > 0$ in such particular case. The results of HDSmP and other probability transformations are listed in Table III.

Table III
EXPERIMENTAL RESULTS FOR EXAMPLE 3.

Approaches	Propositions			E_H
	θ_1	θ_2	θ_3	
BetP	0.1000	0.1000	0.8000	0.5871
PrPl	0.0562	0.0562	0.8876	0.3911
PrBel	NaN	NaN	NaN	NaN
DSmP_0	0.0000	0.0000	1.0000	0.0000
DSmP_0.001	0.0004	0.0004	0.0092	0.0065
HDSmP_0	NaN	NaN	NaN	NaN
HDSmP_0.001	0.0503	0.0503	0.8994	0.3606

It can be seen in Table III that the normalized entropy E_H of HDSmP is relatively small but not the smallest among all the probability transformations used. Naturally, and as already pointed out, $HDSmP_{\epsilon=0}$ cannot be computed in such example because of division by zero. But with the use of the parameter $\epsilon = 0.001$, the mass of $m(\theta_1 \cup \theta_2 \cup \theta_3)$ becomes equally divided and redistributed to the focal elements with cardinality of 2. This justify the necessity of the use of parameter $\epsilon > 0$ in some particular cases when there exist masses equal to zero.

D. Example 4 (vacuous bba)

Let's consider the following particular bba, called the vacuous bba since it represents a fully ignorant source of evidence:

$$\begin{aligned} &m(\theta_1) = 0, \quad m(\theta_2) = 0, \quad m(\theta_3) = 0, \\ &m(\theta_1 \cup \theta_2) = 0, \quad m(\theta_1 \cup \theta_3) = 0, \\ &m(\theta_2 \cup \theta_3) = 0, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 1 \end{aligned}$$

In this example, the mass assignments for all the focal elements with cardinality less than 3 equal to zero. For HDSmP, when $\epsilon > 0$, $m(\theta_1 \cup \theta_2 \cup \theta_3)$ will be divided equally and redistributed to $\{\theta_1 \cup \theta_2\}$, $\{\theta_1 \cup \theta_3\}$ and $\{\theta_2 \cup \theta_3\}$. Similarly, the mass assignments for focal elements with cardinality of 2 (partial ignorances) obtained at the intermediate step will be divided equally and redistributed to singletons included in them. This redistribution is possible for the existence of $\epsilon > 0$ in HDSmP formulas. HDSmP cannot be applied and computed

in such example if one takes $\epsilon = 0$, and that is why one needs to use $\epsilon > 0$ here. The results of HDSmP and other probability transformations are listed in Table IV.

Table IV
EXPERIMENTAL RESULTS FOR EXAMPLE 4.

Approaches	Propositions			E_H
	θ_1	θ_2	θ_3	
BetP	0.3333	0.3333	0.3333	1.0000
PrPl	0.3333	0.3333	0.3333	1.0000
PrBel	NaN	NaN	NaN	NaN
DSmP$_0$	NaN	NaN	NaN	NaN
DSmP$_{0.001}$	0.3333	0.3333	0.3333	1.0000
HDSmP$_0$	NaN	NaN	NaN	NaN
HDSmP$_{0.001}$	0.3333	0.3333	0.3333	1.0000

It can be seen in Tables I – IV that the normalized entropy E_H of HDSmP is always moderate among the other probability transformations it is compared with, and it is normal to get an entropy value with HDSmP bigger than with DSmP because of dilution of uncertainty through the procedure of HDSmP. We have already shown that the entropy criteria is not enough in fact to evaluate the quality a probability transformation [14], and always a compromise must be found between entropy level and numerical robustness of the transformation. Although the entropy should be as small as possible for decision-making, exaggerate small entropy is not always preferred. Because of the way the mass of (partial) ignorances is proportionally redistributed, it is clear that if the mass assignment for a singleton equals to zero in the original bba, then after applying DSmP or HDSmP transformations this mass is unchanged and is kept to zero. This behavior may appear a bit intuitively surprising at the first glance specially if some masses of partial ignorances including this singleton are not equal to zero. This behavior is however normal in the spirit of proportional redistribution because one wants to reduce the PIC value so that if one has no strong support (belief) in a singleton in the original bba, we expect also to have no strong support in this singleton after the transformation is applied which makes perfectly sense. Of course if such behavior is considered as too optimistic or not acceptable because it appears too risky in some applications, it is always possible to choose another transformation instead. The final choice is always left in the hands of the user, or the fusion system designer.

V. CONCLUSIONS

Probability transformation is very crucial for decision-making in evidence theory. In this paper a novel interesting and useful hierarchical probability transformation approach called HDSmP has been proposed, and HDSmP always provides a moderate value of entropy which is necessary for an easier and reliable decision-making support. Unfortunately the PIC (or entropy) level is not the unique useful criterion to evaluate the quality of a probability transformation in general. At least the numerical robustness of the method is also important and must be considered seriously as already shown in our previous works. Therefore, to evaluate any probability transformation more efficiently and to outperform existing transformations (including DSmP and HDSmP) a more general comprehensive evaluation criteria need to be found. The search for such a criteria is under investigations.

ACKNOWLEDGEMENT

This work is supported by National Natural Science Foundation of China (Grant No.61104214 and No. 67114022), Fundamental Research Funds for the Central Universities, China Postdoctoral Science Foundation (No.20100481337, No.201104670) and Research Fund of Shaanxi Key Laboratory of Electronic Information System Integration (No.201101Y17).

REFERENCES

[1] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University, Princeton, 1976.

[2] P. Smets, R. Kennes, *The transferable belief model*, Artificial Intelligence, Vol 66, No. 2, pp. 191-234, 1994.

[3] F. Smarandache, J. Dezert (Editors), *Applications and Advances of DSmT for Information Fusion*, Vol 3, American Research Press, 2009. http://www.gallup.unm.edu/˜smarandache/DSmT-book3.pdf.

[4] W.H. Lv, *Decision-making rules based on belief interval with D-S evidence theory*, book section in Fuzzy Information and Engineering, pp. 619-627, 2007.

[5] J. Sudano, *Pignistic probability transforms for mixes of low- and high-probability events*, in Proc. of Int. Conf. on Information Fusion 2001, Montreal, Canada, , TUB3, pp. 23–27, August, 2001.

[6] J. Sudano, *The system probability information content (PIC) relationship to contributing components, combining independent multisource beliefs, hybrid and pedigree pignistic probabilities*, in Proc. of Int. Conf. on Information Fusion 2002, Annapolis, USA, Vol 2, pp. 1277–1283, July 2002.

[7] J. Sudano, *Equivalence between belief theories and naive Bayesian fusion for systems with independent evidential data - Part I, The theory*, in Proc. of Int. Conf. on Information Fusion 2003, Cairns, Australia, Vol. 2, pp. 1239–1243, July 2003.

[8] J. Sudano, *Yet another paradigm illustrating evidence fusion (YAPIEF)*, in Proc. of Int. Conf. on Information Fusion 2006, Florence, pp. 1–7, July 2006.

[9] J. Dezert, F. Smarandache, *A new probabilistic transformation of belief mass assignment*, in Int. Conf. on Information Fusion 2008, Germany, Cologne, pp. 1-8, June 30^{th} – July 3rd, 2008.

[10] J. Sudano, *Belief fusion, pignistic probabilities, and information content in fusing tracking attributes*, in Proc. of Radar Conference, Volume, Issue, 26-29 , pp. 218 – 224, 2004.

[11] Y. Deng, W. Jiang, Q. Li, *Probability transformation in transferable belief model based on fractal theory* (in Chinese), in Proc. of Chinese Conf. on Information Fusion 2009, Yantai, China, pp. 10-13, 79, Nov. 2009

[12] W. Pan, H.J. Yang, *New methods of transforming belief functions to pignistic probability functions in evidence theory*, Int. Workshop on Intelligent Systems and Applications, pp. 1–5, Wuhan, China, May, 2009.

[13] F. Cuzzolin, *On the properties of the Intersection probability*, submitted to the Annals of Mathematics and AI, Feb. 2007.

[14] D.Q. Han, J. Dezert, C.Z. Han, Y. Yang, *Is Entropy Enough to Evaluate the Probability Transformation Approach of Belief Function?*, in Proc. of 13th Int. Conf. on Information Fusion Conference, pp. 1–7, Edinburgh, UK, July 26-29th, 2010.

[15] P. Smets, *Decision making in the TBM: the necessity of the pignistic transformation*, in Int. Journal of Approximate Reasoning, Vol 38, pp. 133–147, 2005.

[16] J. Dezert, D.Q. Han, Z.G. Liu, J.-M. Tacnet, *Hierarchical proportional redistribution for bba approximation*, in Proceedings of the 2nd International Conference on Belief Functions, Compiègne, France, May 9-11, 2012.

[17] J. Dezert, D.Q. Han, Z.G. Liu, J.-M. Tacnet, *Hierarchical Proportional Redistribution principle for uncertainty reduction and bba approximation*, in Proceedings of the 10th World Congress on Intelligent Control and Automation (WCICA 2012), Beijing, China, July 6-8, 2012.

Neutrosophic Masses & Indeterminate Models

Applications to Information Fusion

Florentin Smarandache

Originally published as Smarandache F., *Neutrosophic Masses & Indeterminate Models*, in Proc. of Fusion 2012, Singapore, July 2012, and reprinted with permission.

***Abstract*—In this paper we introduce the *indeterminate models* in information fusion, which are due either to the existence of some *indeterminate elements* in the fusion space or to some *indeterminate masses*. The best approach for dealing with such models is the neutrosophic logic.**

Keywords: neutrosophic logic; indeterminacy; indeterminate model; indeterminate element; indeterminate mass; indeterminate fusion rules; DSmT; DST; TBM;

I. INTRODUCTION

In this paper we introduce for the first time the notions of indeterminate mass (bba), indeterminate element, indeterminate intersection, and so on. We give an example of neutrosophic dynamic fusion using two classical masses, defined on a determinate frame of discernment, but having indeterminate intersections in the super-power set S^{Θ} (the fusion space). We also adjust several classical fusion rules (*PCR5* and *DSmH*) to work for indeterminate intersections instead of empty intersections.

References [3]-[13] show a wide variety of applications of the neutrosophic logic and set, based on indeterminacy, in information technology.

Let Θ be a frame of discernment, defined as:

$$\Theta = \{\phi_1, \phi_2, \ldots, \phi_n\}, n \geq 2, \quad (1)$$

and its Super-Power Set (or fusion space):

$$S^{\Theta} = (\Theta, \cup, \cap, \complement) \quad (2)$$

which means the set Θ closed under union, intersection, and respectively complement.

This paper is organized as follows: we present the neutrosophic logic, the indeterminate masses, elements and models, and give an example of indeterminate intersection.

II. INDETERMINATE MASS

A. Neutrosophic Logic

Neutrosophic Logic (NL) [1] started in 1995 as a generalization of the fuzzy logic, especially of the intuitionistic fuzzy logic. A logical proposition P is characterized by three neutrosophic components:

$$NL(P) = (T, I, F) \quad (3)$$

where T is the degree of truth, F the degree of falsehood, and I the degree of indeterminacy (or neutral, where the name "neutro-sophic" comes from, i.e. neither truth nor falsehood but in between – or included-middle principle), and with:

$$T, I, F \subseteq]^{-}0,1^{+}[\quad (4)$$

where $]^{-}0,1^{+}[$ is a non-standard interval.

In this paper, for technical proposal, we can reduce this interval to the standard interval $[0, 1]$.

The main distinction between neutrosophic logic and intuitionistic fuzzy logic (IFL) is that in NL the sum $T+I+F$ of the components, when T, I, and F are crisp numbers, does not need to necessarily be 1 as in IFL, but it can also be less than 1 (for incomplete/missing information), equal to 1 (for complete information), or greater than 1 (for paraconsistent/contradictory information).

The combination of neutrosophic propositions is done using the neutrosophic operators (especially $\wedge$, $\vee$).

B. Neutrosophic Mass

We recall that a classical mass $m(.)$ is defined as:

$$m : S^{\Theta} \rightarrow [0,1] \quad (5)$$

such that

$$\sum_{X \in S^{\Theta}} m(X) = 1 \quad (6)$$

We extend this classical basic belief assignment (mass) *m(.)* to a neutrosophic basic belief assignment (nbba) (or neutrosophic mass) $m_n(.)$ in the following way.

$$m_n : S^\Theta \rightarrow [0,1]^3 \quad (7)$$

with

$$m_n(A) = (T(A), I(A), F(A)) \quad (8)$$

where *T(A)* means the (local) chance that hypothesis *A* occurs, *F(A)* means the (local) chance that hypothesis *A* does not occur (nonchance), while *I(A)* means the (local) indeterminate chance of *A* (i.e. knowing neither if *A* occurs nor if *A* doesn't occur),

such that:

$$\sum_{X \in S^\Theta} [T(X) + I(X) + F(X)] = 1. \quad (9)$$

In a more general way, the summation (9) can be less than 1 (for incomplete neutrosophic information), equal to 1 (for complete neutrosophic information), or greater than 1 (for paraconsistent/conflicting neutrosophic information). But in this paper we only present the case when summation (9) is equal to 1.

Of course,

$$0 \leq T(A), I(A), F(A) \leq 1. \quad (10)$$

A *basic belief assignment* (or *mass*) is considered *indeterminate* if there exist at least an element $A \in S^\Theta$ such that *I(A) > 0*, i.e. there exists some indeterminacy in the chance of at least an element *A* for occurring or for not occurring. Therefore, a neutrosophic mass which has at least one element *A* with *I(A) > 0* is an indeterminate mass.

A classical mass *m(.)* as defined in equations (5) and (6) can be extended under the form of a neutrosophic mass $m_n'(.)$ in the following way:

$$m_n' : S^\Theta \rightarrow [0,1]^3 \quad (11)$$

with

$$m_n'(A) = (m(A), 0, 0) \quad (12)$$

but reciprocally it does not work since *I(A)* has no correspondence in the definition of the classical mass.

We just have *T(A) = m(A)* and *F(A) = m(C(A))*, where *C(A)* is the complement of *A*. The non-null *I(A)* can, for example, be roughly approximated by the total ignorance mass $m(\Theta)$, or better by the partial ignorance mass $m(\Theta_I)$ where Θ_I is the union of all singletons that have some non-zero indeterminacy, but these mean less accuracy and less refinement in the fusion.

If *I(X) = 0* for all $X \in S^\Theta$, then the neutrosophic mass is simply reduced to a classical mass.

III. INDETERMINATE ELEMENT

We have two types of elements in the fusion space S^Θ, *determinate elements* (which are well-defined), and *indeterminate elements* (which are not well-defined; for example: a geographical area whose frontiers are vague; or let's say in a murder case there are two suspects, *John* – who is known/determinate element – but he acted together with another man *X* (since the information source saw *John* together with an unknown/unidentified person) – therefore *X* is an indeterminate element).

Herein we gave examples of singletons as indeterminate elements just in the frame of discernment Θ, but indeterminate elements can also result from the combinations (unions, intersections, and/or complements) of determinate elements that form the super-power set S^Θ. For example, *A* and *B* can be determinate singletons (we call the elements in Θ as singletons), but their intersection $A \cap B$ can be an indeterminate (unknown) element, in the sense that we might not know if $A \cap B = \phi$ or $A \cap B \neq \phi$.

Or *A* can be a determinate element, but its complement *C(A)* can be indeterminate element (not well-known), and similarly for determinate elements *A* and *B*, but their $A \cup B$ might be indeterminate.

Indeterminate elements in S^Θ can, of course, result from the combination of indeterminate singletons too. All depends on the problem that is studied.

A frame of discernment which has at least an indeterminate element is called *indeterminate frame of discernment*. Otherwise, it is called *determinate frame of discernment*. Similarly we call an *indeterminate fusion space* (S^Θ) that fusion space which has at least one indeterminate element. Of course an indeterminate frame of discernment spans an indeterminate fusion space.

An *indeterminate source of information* is a source which provides an indeterminate mass or an indeterminate fusion space. Otherwise it is called a *determinate source of information*.

IV. INDETERMINATE MODEL

An *indeterminate model* is a model whose fusion space is indeterminate, or a mass that characterizes it is indeterminate.

Such case has not been studied in the information fusion literature so far. In the next sections we'll present some examples of indeterminate models.

V. CLASSIFICATION OF MODELS

In the classical fusion theories all elements are considered determinate in the Closed World, except in Smets' Open World where there is some room (i.e. mass assigned to the empty set) for a possible unknown missing singleton in the frame of discernment. So, the Open World has a probable indeterminate element, and thus its frame of discernment is indeterminate. While the Closed World frame of discernment is determinate.

In the Closed World in Dezert-Smarandache Theory there are three models classified upon the types of singleton intersections: Shafer's Model (where all intersections are empty), Hybrid Model (where some intersections are empty, while others are non-empty), and Free Model (where all intersections are non-empty).

We now introduce a fourth category, called *Indeterminate Model* (where at least one intersection is indeterminate/unknown, and in general at least one element of the fusion space is indeterminate). We do this because in practical problems we don't always know if an intersection is empty or nonempty. As we still have to solve the problem in the real time, we have to work with what we have, i.e. with indeterminate models.

The *indeterminate intersection* cannot be refined (because not knowing if $A \cap B$ is empty or nonempty, we'd get two different refinements: $\{A, B\}$ when intersection is empty, and $\{A\backslash B, B\backslash A, A \cap B\}$ when intersection is nonempty).

The *percentage of indeterminacy* of a model depends on the number of indeterminate elements and indeterminate masses.

By default: the sources, the masses, the elements, the frames of discernment, the fusion spaces, and the models are supposed determinate.

VI. AN EXAMPLE OF INFORMATION FUSION WITH AN INDETERMINATE MODEL

We present the below example.

Suppose we have two sources, $m_1(.)$ and $m_2(.)$, such that:

	A	B	C	$A \cup B \cup C$	$A \cap B$ = *Ind.*	$A \cap C$ = ϕ	$B \cap C$ = *Ind.*
m_1	0.4	0.2	0.3	0.1			
m_2	0.1	0.3	0.2	0.4			
m_{12}	.21	.17	.20	.04	.14	.11	.13

Table 1

Applying the conjunction rule to m_1 and m_2 we get $m_{12}(.)$ as shown in Table 1.

The frame of discernment is $\Theta = \{A, B, C\}$. We know that $A \cap C$ is empty, but we don't know the other two intersections: we note them as $A \cap B = ind.$ and $B \cap C = ind,.$ where *ind.* means indeterminate.

Using the Conjunctive Rule to fusion m_1 and m_2, we get $m_{12}(.)$:

$$\forall A \in S^{\Theta} \setminus \phi, m_{12}(A) = \sum_{\substack{X,Y \in S^{\Theta} \\ A=X \cap Y}} m_1(X) m_2(Y) . \quad (13)$$

Whence: $m_{12}(A)=0.21$, $m_{12}(B)=0.17$, $m_{12}(C)=0.20$, $m_{12}(A \cup B \cup C)=0.04$, and for the intersections:

$m_{12}(A \cap B)=0.14$, $m_{12}(A \cap C)=0.11$, $m_{12}(B \cap C)=0.13$.

We then use the *PCR5* fusion rule style to redistribute the masses of these three intersections. We recall *PCR5* for two sources:

(14)

$$\forall A \in S^{\Theta} \setminus \phi,$$

$$m_{12\,PCR5}(A) = m_{12}(A) + \sum_{\substack{X \in S^{\Theta} \setminus \{\phi\} \\ X \cap A = \phi}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A) + m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A) + m_1(X)} \right]$$

a) $m_{12}(A \cap C)=0.11$ is redistributed back to A and C because $A \cap C = \phi$, according to the PCR5 style.

Let $\alpha 1$ and $\alpha 2$ be the parts of mass 0.11 redistributed back to A, and $\gamma 1$ and $\gamma 2$ be the parts of mass 0.11 redistributed back to C.

We have the following proportionalizations:

$$\frac{\alpha 1}{0.4} = \frac{\gamma 1}{0.2} = \frac{0.4 \cdot 0.2}{0.4 + 0.2} = 0.133333,$$

whence $\alpha 1 = 0.4(0.133333) \approx 0.053333$
and $\gamma 1 = 0.2(0.13333) \approx 0.026667$.
Similarly:

$$\frac{\alpha 2}{0.1} = \frac{\gamma 2}{0.3} = \frac{0.1 \cdot 0.3}{0.1 + 0.3} = 0.075,$$

whence $\alpha 2 = 0.1(0.075) = 0.0075$
and $\gamma 2 = 0.3(0.075) = 0.0225$.
Therefore the mass of A, which can also be noted as $T(A)$ in a neutrosophic mass form, receives from 0.11 back:
$\alpha 1 + \alpha 2 = 0.053333 + 0.0075 = 0.060833$,
while the mass of C, or $T(C)$ in a neutrosophic form, receives from 0.11 back:
$\gamma 1 + \gamma 2 = 0.026667 + 0.0225 = 0.049167$.
We verify our calculations: $0.060833 + 0.049167 = 0.11$.

b) $m_{12}(A \cap B)=0.14$ is redistributed back to the indeterminate parts of the masses of A and B respectively, namely $I(A)$ and $I(B)$ as noted in the neutrosophic mass form, because $A \cap B = Ind.$ We follow the same PCR5 style as done in classical PCR5 for empty intersections (as above).

Let $\alpha 3$ and $\alpha 4$ be the parts of mass 0.14 redistributed back to $I(A)$, and $\beta 1$ and $\beta 2$ be the parts of mass 0.14 redistributed back to $I(B)$.

We have the following proportionalizations:

$$\frac{\alpha 3}{0.4} = \frac{\beta 1}{0.3} = \frac{0.4 \cdot 0.3}{0.4 + 0.3} = 0.171429,$$

whence $\alpha 3 = 0.4(0.171429) \approx 0.068572$
and $\beta 1 = 0.3(0.171429) \approx 0.051428$.
Similarly:

$$\frac{\alpha 4}{0.1} = \frac{\beta 2}{0.2} = \frac{0.1 \cdot 0.2}{0.1 + 0.2} = 0.066667$$

whence $\alpha 4 = 0.1(0.066667) \approx 0.006667$
and $\beta 2 = 0.2(0.066667) \approx 0.013333$.
Therefore, the indeterminate mass of A, $I(A)$ receives from 0.14 back:

$\alpha 3 + \alpha 4 = 0.068572 + 0.006667 = 0.075239$

and the indeterminate mass of *B, I(B),* receives from 0.14 back:

$\beta 1+\beta 2 = 0.051428+0.013333=0.064761$.

c) Analougously, $m_{12}(B\cap C)=0.13$ is redistributed back to the indeterminate parts of the masses of *B* and *C* respectively, namely *I(B)* and *I(C)* as noted in the neutrosophic mass form, because $B\cap C=Ind.$ also following the PCR5 style. Whence *I(B)* gets back 0.065 and *I(C)* also gets back 0.065.

Finally we sum all results obtained from firstly using the Conjunctive Rule [Table 1] and secondly redistributing the intersections masses with PCR5 [sections a), b), and c) from above]:

	T(A)	T(B)	T(C)	$T(\Theta)$	I(A)	I(B)	I(C)
m_{12}	.21	.17	.20	.04			
addi-tions	.0075 .053 333		.022 5 .026 667		.068 572 .006 667	.051 428 .013 333 .02 .045	.04 .045
$m_{12PCR5I}$	.270 833	.17	.249 167	.04	.075 239	.129 761	.065

Table 2

where $\Theta = A \cup B \cup C$ is the total ignorance.

VII. BELIEF, DISBELIEF, AND UNCERTAINTY

In classical fusion theory there exist the following functions:

Belief in *A* with respect to the bba *m(.)* is:

$$Bel(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\subseteq A}} m(X) \tag{15}$$

Disbelief in *A* with respect to the bba *m(.)* is:

$$Dis(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A=\phi}} m(X) \tag{16}$$

Uncertainty in *A* with respect to the bba *m(.)* is:

$$U(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A\neq\phi \\ X\cap C(A)\neq\phi}} m(X), \tag{17}$$

where *C(A)* is the complement of A with respect to the total ignorance Θ.

Plausability of *A* with respect to the bba *m(.)* is:

$$Pl(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A\neq\phi}} m(X) \tag{18}$$

VIII. NEUTROSOPHIC BELIEF, NEUTROSOPHIC DISBELIEF, AND NEUTROSOPHIC UNDECIDABILITY

Let's consider a neutrosophic mass $m_n(.)$ as defined in formulas (7) and (8), $m_n(X) = (T(X), I(X), F(X))$ for all $X \in S^{\Theta}$.

We extend formulas (15)-(18) from *m(.)* to $m_n(.)$:

Neutrosophic Belief in *A* with respect to the nbba $m_n(.)$ is:

$$NeutBel(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\subseteq A}} T(X) + \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A=\phi}} F(X) \tag{19}$$

Neutrosophic Disbelief in *A* with respect to the nbba $m_n(.)$ is:

$$NeutDis(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A=\phi}} T(X) + \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\subseteq A}} F(X) \tag{20}$$

Neutrosophic Uncertainty in *A* with respect to the nbba $m_n(.)$ is

$$\begin{aligned} NeutU(A) &= \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A\neq\phi \\ X\cap C(A)\neq\phi}} T(X) + \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A\neq\phi \\ X\cap C(A)\neq\phi}} F(X) \\ &= \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A\neq\phi \\ X\cap C(A)\neq\phi}} [T(X)+F(X)] \end{aligned} \tag{21}$$

We now introduce the **Neutrosophic Global Indeterminacy in *A*** with respect to the nbba $m_n(.)$ as a sum of local indeterminacies of the elements included in *A*:

$$NeutGlobInd(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\subseteq A}} I(X) \tag{22}$$

And afterwards we define another function called **Neutrosophic Undecidability about *A*** with respect to the nbba $m_n(.)$:

$$NeutUnd(A) = NeutU(A) + NeutGlobInd(A) \tag{23}$$

or

$$NeutUnd(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A\neq\phi \\ X\cap C(A)\neq\phi}} [T(X)+F(X)] + \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\subseteq A}} I(X) \tag{24}$$

Neutrosophic Plausability of A with respect to the nbba $m_n(.)$ is:

$$NeutPl(A) = \sum_{\substack{X\in S^{\Theta}\setminus\{\phi\} \\ X\cap A\neq\phi}} T(X) + \sum_{\substack{Y\in S^{\Theta}\setminus\{\phi\} \\ C(Y)\cap A\neq\phi}} F(Y) \tag{25}$$

In the previous example let's compute *NeutBel(.), NeutDis(.),* and *NeutUnd(.)*:

	A	*B*	*C*	$A \cup B \cup C$
NeutBel	*0.270833*	*0.17*	*0.249167*	*0.73*
NeutDis	*0.419167*	*0.52*	*0.440833*	*0*
NeutGlobInd	*0.115239*	*0.169761*	*0.105*	*0*
Total	*0.805239* $\neq$ *1*	*0.859761* $\neq$ *1*	*0.795* $\neq$ *1*	*0.73* $\neq$ *1*

Table 3

As we see, for indeterminate model we cannot use the intuitionistic fuzzy set or intuitionistic fuzzy logic since the sum *NeutBel(X)+NeutDis(X)+NeutGlobInd(X)* is less than 1. In this case we use the neutrosophic set or logic which can deal with incomplete information.

The sum is less than 1 because there is missing information (we don't know if some intersections are empty or not).

For example:

NeutBel(A)+NeutDis(A)+NeutGlobInd(A)=0.805239 =1-I(B)-I(C).

Similarly,

NeutBel(B)+NeutDis(B)+NeutGlobInd(B)=0.859761 =1-I(A)-I(C).

NeutBel(C)+NeutDis(C)+NeutGlobInd(C)=0.795 =1-I(A)-I(B)

and

NeutBel($A \cup B \cup C$*)+NeutDis(*$A \cup B \cup C$*) +NeutGlobInd(*$A \cup B \cup C$*)=0.73=1-I(A)-I(B)-I(C).*

IX. NEUTROSOPHIC DYNAMIC FUSION

A Neutrosophic Dynamic Fusion is a dynamic fusion where some indeterminacy occurs: with respect to the mass or with respect to some elements.

The solution of the above indeterminate model which has missing information, using the neutrosophic set, is consistent in the classical dynamic fusion in the case we receive part (or total) of the missing information.

In the above example, let's say we find out later in the fusion process that $A \cap B = \phi$. That means that the mass of indeterminacy of *A, I(A)=0.075239*, is transferred to *A*, and the masses of indeterminacy of *B* (resulted from $A \cap B$ only) - i.e. *0.051428* and *0.13333* - are transferred to *B*. We get:

	A	B	C	Θ	I(A)	I(B)	I(C)	A⊠B	A⊠C
m	.270833	.17	.249167	.04	0	.065	.065	0	0
+	.075239	.051428 .013333							
m_N	.346072	.234761	.249167	.04	0	.065	.065	0	0

Table 4

where $\Theta = A \cup B \cup C$ is the total ignorance.

The sum *NeutBel(X)+NeutDis(X)+NeutBlogInd(X)* increases towards 1, as indeterminacy *I(X)* decreases towards 0, and reciprocally.

When we have complete information we get *NeutBel(X)+NeutDis(X)+NeutGlobInd(X)=1* and in this case we have an intuitionistic fuzzy set, which is a particular case of the neutrosophic set.

Let's suppose once more, considering the neutrosophic dynamic fusion, that afterwards we find out that $B \cap C \neq \phi$. Then, from *Table 4* the masses of indeterminacies of *B, I(B)* (0.065 = 0.02 + 0.045, resulted from $B \cap C$ which was considered indeterminate at the beginning of the neutrosophic dynamic fusion), and that of *C, I(C)=0.065*, go now to $B \cap C$. Thus, we get:

	A	B	C	Θ	I(A)	I(B)	I(C)	A⊠B	A⊠C	B⊠C
m_N	.346072	.234761	.249167	.04	0	.065	.065	0	0	0
-/+						-.065	-.065			+.065 +.065
m_{NN}	.346072	.234761	.249167	.04	0	0	0	0	0	.13

Table 5

X. MORE REDISTRIBUTION VERSIONS FOR INDETERMINATE INTERSECTIONS OF DETERMINATE ELEMENTS

Besides *PCR5*, it is also possible to employ other fusion rules for the redistribution, such as follows:

a. For the masses of the empty intersections we can use *PCR1-PCR4, URR, PURR, Dempster's Rule*, etc. (in general any fusion rule that first uses the conjunctive rule, and then a redistribution of the masses of empty intersections).
b. For the masses of the indeterminate intersections we can use *DSm Hybrid (DSmH)* rule to transfer the mass $m_{12}(X \cap Y = ind.)$ to $X \cup Y$, since $X \cup Y$ is a kind of uncertainty related to X, Y. In our opinion, a better approach in this case would be to redistributing the empty intersection masses using the *PCR5* and the indeterminate intersection masses using the *DSmH*, so we can combine two fusion rules into one:

Let $m_1(.)$ and $m_2(.)$ be two masses. Then:

$$m_{12\,PCR5\ /\ DSmH}(A) = \sum_{\substack{X,Y\in S^\Theta\setminus\{\phi\}\\ X\cap Y=A}} m_1(X)m_2(Y)$$

$$+ \sum_{\substack{X\in S^\Theta\setminus\{\phi\}\\ X\cap A=\phi}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A)+m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A)+m_1(X)}\right]$$

$$+ \sum_{\substack{X,Y\in S^\Theta\setminus\{\phi\}\\ X\cap Y=ind.\\ X\cup Y=A}} m_1(X)m_2(Y)$$

$$= \sum_{\substack{X,Y\in S^\Theta\setminus\{\phi\}\\ \{X\cap Y=A\}\vee\{(X\cap Y=ind.)\wedge(X\cup Y=A)\}}} m_1(X)m_2(Y)$$

$$+ \sum_{\substack{X\in S^\Theta\setminus\{\phi\}\\ X\cap A=\phi}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A)+m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A)+m_1(X)}\right] \quad (26)$$

Yet, the best approach, for an indeterminate intersection resulted from the combination of two classical masses $m_1(.)$ and $m_2(.)$ defined on a determinate frame of discernment, is the first one:

- Use the *PCR5* to combine the two sources: formula (14).
- Use the *PCR5-ind* [adjusted from classical *PCR5* formula (14)] in order to compute the indeterminacies of each element involved in indeterminate intersections :

$$\forall A\in S^\Theta\setminus\phi,$$

$$m_{12\,PCR5\,Ind}(I(A)) = \sum_{\substack{X\in S^\Theta\setminus\{\phi\}\\ X\cap A=ind.}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A)+m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A)+m_1(X)}\right] \quad (27)$$

- Compute *NeutBel(.), NeutDis(.), NeutGlobInd(.)* of each element.

CONCLUSION

In this paper we introduced for the first time the notions of indeterminate mass (bba), indeterminate element, indeterminate intersection, and so on. We gave an example of neutrosophic dynamic fusion using two classical masses, defined on a determinate frame of discernment, but having indeterminate intersections in the super-power set S^Θ (the fusion space). We adjusted several classical fusion rules (*PCR5* and *DSmH*) to work for indeterminate intersections instead of empty intersections.

Then we extended the classical *Bel(.), Dis(.)* {also called *Dou(.)*, i.e Dough} and the uncertainty *U(.)* functions to their respectively neutrosophic correspondent functions that use the neutrosophic masses, i.e. to the *NeutBel(.), NeutDis(.), NeutU(.)* and to the undecidability function *NeutUnd(.)* . We have also introduced the Neutrosophic Global Indeterminacy function, *NeutGlobInd(.)*, which together with *NeutU(.)* form the Neut*Und(.)* function.

In our first example the mass of $A\cap B$ is determined (it is equal to *0.14*), but the element $A\cap B$ is indeterminate (we don't know if it empty or not).

But there are cases when the element is determinate (let's say a suspect John), but its mass could be indeterminate as given by a source of information {for example $m_n(John) = (0.4, 0.1, 0.2)$, i.e. there is some mass indeterminacy: $I(John) = 0.2 > 0$}.

These are the distinctions between the indeterminacy of an element, and the indeterminacy of a mass.

ACKNOWLEDGMENT

The author would like to thank Dr. Jean Dezert for his comments related to the indeterminate models from our Fall 2011 email correspondence.

REFERENCES

[1] F. Smarandache, *A Unifying Field in Logics: Neutrosophic Logic*, Multiple Valued Logic / An International Journal, Vol. 8, No. 3, 2002, pp. 385-438.

[2] F. Smarandache, J. Dezert (editors), *Advances and Applications of DSmT for Information Fusion*, Am. Res. Press, Rehoboth,2004-2009, http://fs.gallup.unm.edu/DsmT.htm.

[3] F. G. Lupiáñez: *Interval neutrosophic sets and Topology*, Kybernetes **38** (2009), 621-624.

[4] Andrew Schumann, *Neutrosophic logics on Non-Archimedean Structures*, Critical Review, Creighton University, USA, Vol. III, 36-58, 2009.

[5] Fu Yuhua, Fu Anjie, Zhao Ge., *Positive, Negative and Neutral Law of Universal Gravitation*, New Science and Technology, 2009 (12), 30-32.

[6] Monoranjan Bhowmik and Madhumangal Pal, *Intuitionistic Neutrosophic Set*, Journal of Information and Computing Science, England, Vol. 4, No. 2, 2009, pp. 142-152.

[7] Wen Ju and H. D. Cheng, *Discrimination of Outer Membrane Proteins using Reformulated Support Vector Machine based on Neutrosophic Set*, Proceedings of the 11th Joint Conference on Information Sciences (2008), Published by Atlantis Press.

[8] Goutam Bernajee, *Adaptive fuzzy cognitive maps vs neutrosophic cognitive maps: decision support tool for knowledge based institution*, Journal of Scientific and Industrial Research, 665-673, Vol. 67, 2008.

[9] Smita Rajpal, M.N. Doja, Ranjit Biswas, *A Method of Imprecise Query Solving*, International Journal of Computer Science and Network Security, Vol. 8 No. 6, pp. 133-139, June 2008, http://paper.ijcsns.org/07_book/200806/20080618.pdf

[10] Jose L. Salmeron, F. Smarandache, *Redesigning Decision Matrix Method with an indeterminacy-based inference process*, Advances in Fuzzy Sets and Systems, Vol. 1(2), 263-271, 2006.

[11] P. Kraipeerapun, C. C. Fung, W. Brown and K. W. Wong, , *Neural network ensembles using interval neutrosophic sets and bagging for mineral prospectivity prediction and quantification of uncertainty*, 2006 IEEE Conference on Cybernetics and Intelligent Systems, 7-9 June 2006, Bangkok, Thailand.

[12] Goutam Bernajee, *Adaptive fuzzy cognitive maps vs neutrosophic cognitive maps: decision support tool for knowledge based institution*, Journal of Scientific and Industrial Research, 665-673, Vol. 67, 2008.

[13] Anne-Laure Jousselme, Patrick Maupin, *Neutrosophy in situation analysis*, Proceedings of Fusion 2004 Int. Conf. on Information Fusion, pp. 400-406, Stockholm, Sweden, June 28-July 1, 2004.

Extended PCR Rules for Dynamic Frames

Florentin Smarandache
Jean Dezert

Originally published as Smarandache F., Dezert J., *Extended PCR Rules for Dynamic Frames*, Proc. Of Fusion 2012, Singapore, July 2012, and reprinted with permission.

***Abstract*—In most of classical fusion problems modeled from belief functions, the frame of discernment is considered as static. This means that the set of elements in the frame and the underlying integrity constraints of the frame are fixed forever and they do not change with time. In some applications, like in target tracking for example, the use of such invariant frame is not very appropriate because it can truly change with time. So it is necessary to adapt the Proportional Conflict Redistribution fusion rules (PCR5 and PCR6) for working with dynamical frames. In this paper, we propose an extension of PCR5 and PCR6 rules for working in a frame having some non-existential integrity constraints. Such constraints on the frame can arise in tracking applications by the destruction of targets for example. We show through very simple examples how these new rules can be used for the belief revision process.**

Keywords: Information fusion, DSmT, integrity constraints, belief functions.

I. Introduction

In most of classical fusion problems using belief functions, the frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ is considered static. This means that the set of elements in the frame (assumed to be non-empty and distinct) and the underlying integrity constraints of the frame[1] are fixed and they do not change with time. In some applications however, like in target tracking and battlefield surveillance for example, the use of such invariant frame is not very appropriate because it can truly change with time depending on the evolution of the events. So it is necessary to adapt the Proportional Conflict Redistribution fusion rules (PCR5 and PCR6) for working with dynamical frames. In this paper, we study in details how to work with PCR5 or PCR6 fusion rules in a dynamical frame subject to non-existential integrity constraint, when one or several elements of the frame disappear. This phenomena can occur in some applications, specially in defense and battlefield surveillance when foe targets (considered as element of the frame) can be shot and entirely destroyed and the initial belief one has on threat assessment must be revised according to the knowledge one has on this new fact obtained from intelligence services or observations systems. We show through very simple examples how this problem can be solved using PCR principle.

Example 1: Let's consider the set of three targets at a given time k to be $\Theta_k = \{\theta_1, \theta_2, \theta_3\}$ with $\theta_i \neq \emptyset$, $i = 1, 2, 3$ and assume that Θ_k satisfies Shafer's model (i.e. the targets are all distinct and exhaustive) and we work with normalized bba's. Suppose one has two basic belief assignments (bba) $m_1(.)$ and $m_2(.)$ defined with respect to the power-set of Θ_k given by two distinct sources of evidence to characterize their beliefs in the most threatening target. Let's assume that one receives at $k+1$ a new information confirming that one target, say target θ_3, has been destroyed. The problem one needs to solve is how to combine efficiently $m_1(.)$ and $m_2(.)$ taking into account this new non-existential integrity constraint $\theta_3 \equiv \emptyset$ in the new model of the frame to establish the most threatening and surviving targets belonging to $\Theta_{k+1} = \{\theta_1, \theta_2\}$.

The contribution of this paper is to propose a solution to such kind of belief revision problem involving dynamical frames including non-existential constraints on some of its elements. This paper is organized as follows. In section 1, we briefly recall the basis of DSmT (Dezert-Smarandache Theory) [4] and its main rule of combination (PCR5 and PCR6) for the fusion of bba's in a static frame. In section 2, we present an improvement/adaptation of PCR rules to work on frames with non-existential constraints (dynamical frames). In section 3, we apply our method on some examples. Conclusions are then given in section 4.

II. Basics of DSmT

The purpose of the development of Dezert-Smarandache Theory (DSmT) [4] is to overcome the limitations of Dempster-Shafer Theory (DST) [3] mainly by proposing new underlying models for the frames of discernment in order to fit better with the nature of real problems, and by proposing new efficient combination and conditioning rules. In DSmT framework, the elements θ_i, $i = 1, 2, \ldots, n$ of a given frame Θ are not necessarily exclusive, and there is no restriction on θ_i but their exhaustivity. The hyper-power set D^Θ in DSmT, the hyper-power set is defined as the set of all composite propositions built from elements of Θ with operators $\cup$ and $\cap$. For instance, if $\Theta = \{\theta_1, \theta_2\}$, then $D^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_1 \cap \theta_2, \theta_1 \cup \theta_2\}$. The hyper-power set D^Θ reduces to classical power-set 2^Θ as soon as we assume exclusivity between the elements of the frame (this is Shafer's model). A (generalized) basic belief assignment (bba for short) is defined as the mapping $m : D^\Theta \to [0, 1]$. The generalized belief and plausibility functions are defined in almost the same manner as in DST. More precisely, from a general frame Θ, we define a map $m(.) : D^\Theta \to [0, 1]$ associated to a given

[1] This is also called the model for Θ which can correspond to DSm free, DSm hybrid or Shafer's models in DSmT framework [4].

body of evidence $\mathcal{B}$ as

$$m(\emptyset) = 0 \qquad \text{and} \qquad \sum_{A \in D^\Theta} m(A) = 1 \tag{1}$$

The quantity $m(A)$ is called the *generalized* basic belief assignment/mass (or just "bba" for short) of A.
The *generalized* credibility and plausibility functions are defined in almost the same manner as within DST, i.e.

$$\text{Bel}(A) = \sum_{\substack{B \subseteq A \\ B \in D^\Theta}} m(B) \quad \text{and} \quad \text{Pl}(A) = \sum_{\substack{B \cap A \neq \emptyset \\ B \in D^\Theta}} m(B) \tag{2}$$

Two models[2] (the free model and hybrid model) in DSmT can be used to define the bba's to combine. In the free DSm model, the sources of evidence are combined without taking into account integrity constraints. When the free DSm model does not hold because the true nature of the fusion problem under consideration, we take into account some known integrity constraints[3] and define bba's to combine using the proper hybrid DSm model. Aside offering the possibility to work with different underlying models (not only Shafer's model as within DST), DSmT offers also new efficient combination rules based on proportional conflict redistribution (PCR rules no 5 and no 6) for combining highly conflicting sources of evidence. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. (see [4], Vol. 2 for full justification and examples): $m_{PCR5}(\emptyset) = 0$ and $\forall X \in D^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1, X_2 \in D^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{X_2 \in D^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)} \right] \tag{3}$$

where all denominators in (3) are different from zero. If a denominator is zero, that fraction is discarded. The properties of PCR5 can be found in [2]. Extension of PCR5 for combining qualitative bba's can be found in [4], Vol. 2 & 3. All propositions/sets are in a canonical form. A variant of PCR5, called PCR6 has been proposed by Martin and Osswald in [4], Vol. 2, for combining $s > 2$ sources. The general formulas for PCR5 and PCR6 rules are given in [4], Vol. 2 also. PCR6 coincides with PCR5 when one combines two sources. The difference between PCR5 and PCR6 lies in the way the proportional conflict redistribution is done as soon as three or more sources are involved in the fusion. From the implementation point of view, PCR6 is much more simple to implement than PCR5. For convenience, very basic (not optimized) Matlab codes of PCR5 and PCR6 fusion rules can be found in [4], [5] and from the toolboxes repository on the web [7]. In DSmT framework, the classical pignistic transformation $BetP(.)$ is replaced by the more effective $DSmP(.)$ transformation to estimate the subjective probabilities of hypotheses for decision-making support once the combination of bba's has been done if compromise attitude is chosen. The max of credibility (pessimistic decision attitude) or max of plausibility (optimistic decision attitude) are also possible depending on the preference of decision maker. This topic is out of the scope of this paper and readers interested in decision-making based on DSmP must refer to [4], Vol.3 freely available on the web.

[2]Actually, Shafer's model, considering all elements of the frame as truly exclusive, can be viewed as a special case of hybrid model.

[3]but non-existential integrity constraints as shown in Example 2.

III. Working with non-existential constraints

In this section we show how this problem can be solved from the classical Shafer's approach and then we show how it can be solved with PCR rules to get more specific results.

A. Shafer's approach

Let's consider a finite and discrete frame $\Theta_k = \{\theta_1, \theta_2, \ldots, \theta_n\}$ satisfying Shafer's model with all $\theta_i \neq \emptyset$ at a given time k, and two bba's $m_{1,k}(.)$ and $m_{2,k}(.)$ provided by two distinct sources of evidences. Each bba is defined in the power set 2^{Θ_k}. Let's assume now that at time $k+1$ extra knowledge is given about the non-existence of some elements of Θ_k. We denote such non-existential constraint as NE (the set of Non Existing elemnts). For example, if $\text{NE}_{k+1} = \{\theta_1\}$ means that actually $\theta_1 = \emptyset$, $\text{NE}_{k+1} = \{\theta_1, \theta_2\}$ means that both $\theta_1 = \emptyset$ and $\theta_2 = \emptyset$, and so on. The new frame of discernment we have to work with is then given by $\Theta_{k+1} = \Theta_k \setminus \text{NE}_{k+1}$. The question is how to combine at time $k+1$ the two original bba's $m_{1,k}(.)$ and $m_{2,k}(.)$ one had in taking into account our knowledge on the revised frame Θ_{k+1} obtained from Θ_k and NE_{k+1} ?

Dempster-Shafer Theory (DST) [3] offers a mathematical tool for answering to this question: Dempster-Shafer belief conditioning rule (DSCR) which consists in combining with Dempster-Shafer's rule the prior bba $m(.)$ with the conditioning bba $m_c(.)$ which is only focused on the conditioning event X, i.e. for which $m_c(X) = 1$. Mathematically, $m_{DS}(.|X)$ is then defined[4] by

$$m_{DS}(.|X) = [m \oplus m_c](.) \tag{4}$$

where $\oplus$ corresponds here to Dempster-Shafer's rule of combination and $m_c(X) = 1$.

For solving this fusion problem under non-existential integrity constraints, three methods are a priori possible based on DSCR:

• **The Fusion-Conditioning approach (FC)**: It consists to combine the sources at first and then apply Dempster-Shafer conditioning rule. This corresponds to the following formula:

$$m_{DS\text{-}FC}(.|\Theta_{k+1}) = [[m_{1,k} \oplus m_{2,k}] \oplus m_{c,k}](.) \tag{5}$$

where $\oplus$ corresponds here to Dempster-Shafer's rule of combination and $m_{c,k}(\Theta_{k+1}) = 1$. Note that $m_{c,k}(.)$ refers to the conditioning bba defined in 2^{Θ_k}.

[4]if $m(.)$ and $m_c(.)$ are not in total contradiction of course.

• **The Conditioning-Fusion approach (CF)**: It consists to apply the DS conditioning to the sources at first and then combine the conditioned bba's with Dempster-Shafer rule. This corresponds to the following formula:

$$m_{DS\text{-}CF}(.|\Theta_{k+1}) = [m_{1,k} \oplus m_{c,k}] \oplus [m_{2,k} \oplus m_{c,k}](.) \quad (6)$$

• **The Global Conditioning approach (GC)**: It consists to combine all the bba's altogether in a single step of fusion. This corresponds to the following formula:

$$m_{DS\text{-}GC}(.|\Theta_{k+1}) = [m_{1,k} \oplus m_{2,k} \oplus m_{c,k}](.) \quad (7)$$

Because of the commutativity and associativity of DS rule and since $[m_c \oplus m_c](.) = m_c(.)$ for any conditioning bba focused on only one specific element X, the three previous methods provide exactly the same results. This makes Shafer's approach very appealing since there is no ambiguity in the choice of the method to apply.

B. Example 1 (continued)

Let's take back the Example 1 and consider the two arbitrary prior bba's given in Table I.

bba's\focal elem.	θ_1	θ_2	θ_3	$\theta_1 \cup \theta_2$
Prior: $m_{1,k}(.)$	0.2	0.4	0.3	0.1
Prior: $m_{2,k}(.)$	0.3	0.1	0.4	0.2
Conditioning: $m_{c,k}(.)$	0	0	0	1
DS-FC: $m_{DS\text{-}FC}(.)$	0.4643	0.4643	0	0.0714
DS-CF: $m_{DS\text{-}CF}(.)$	0.4643	0.4643	0	0.0714
DS-GC: $m_{DS\text{-}GC}(.)$	0.4643	0.4643	0	0.0714

Table I
EXAMPLE 1: RESULTS WITH DS-BASED CONDITIONING.

Because in this example $\Theta_k = \{\theta_1, \theta_2, \theta_3\}$ and $\text{NE}_{k+1} = \{\theta_3\}$ then $\Theta_{k+1} = \{\theta_1, \theta_2\}$ (only targets θ_1 and θ_2 survive) and therefore the conditioning bba $m_{c,k}(.)$ is defined by $m_{c,k}(\theta_1 \cup \theta_2) = 1$. In applying DSCR, one gets with three methods the same following result: $m_{DS\text{-}GC}(.|\Theta_{k+1}) = m_{DS\text{-}CF}(.|\Theta_{k+1}) = m_{DS\text{-}FC}(.|\Theta_{k+1})$ as shown in the last three rows of Table I). This symmetrical result in θ_1 and θ_2 is very surprising since clearly the input bba's are asymmetrical in θ_1 and θ_2 and we don't see any intuitive nor rational justification to consider such DSCR-based behavior as efficient for applications.

• **Direct approach**: Note that this result can be also simply obtained in a direct manner using DS rule for combining $m_{1,k}(.)$ with $m_{2,k}(.)$ and in taking into account the constraint $\theta_3 = \emptyset$ in the DS formula. In this example 1, one gets:

$$\begin{aligned} m_{12}(\theta_1) &= m_{1,k}(\theta_1)m_{2,k}(\theta_1) + m_{1,k}(\theta_1)m_{2,k}(\theta_1 \cup \theta_2) \\ &\quad + m_{2,k}(\theta_1)m_{1,k}(\theta_1 \cup \theta_2) = 0.13 \\ m_{12}(\theta_2) &= m_{1,k}(\theta_2)m_{2,k}(\theta_2) + m_{1,k}(\theta_2)m_{2,k}(\theta_1 \cup \theta_2) \\ &\quad + m_{2,k}(\theta_2)m_{1,k}(\theta_1 \cup \theta_2) = 0.13 \\ m_{12}(\theta_1 \cup \theta_2) &= m_{1,k}(\theta_1 \cup \theta_2)m_{1,k}(\theta_1 \cup \theta_2) = 0.02 \end{aligned}$$

For θ_3, one has $m_{12}(\theta_3) = m_{1,k}(\theta_3)m_{1,k}(\theta_3) = 0.12$. Since actually $\theta_3 = \emptyset$, then $m_{12}(\theta_3 = \emptyset) = 0.12$ must be added to mass already committed to the empty set coming from other possible conflicting conjunctions so that finally one will get the total conflicting mass $m_{12}(\emptyset) = 0.72$. After normalization step, we finally get

$$\begin{aligned} m_{DS}(\theta_1) &= \frac{m_{12}(\theta_1)}{1 - m_{12}(\emptyset)} = 0.13/0.28 = 0.4643 \\ m_{DS}(\theta_2) &= \frac{m_{12}(\theta_2)}{1 - m_{12}(\emptyset)} = 0.13/0.28 = 0.4643 \\ m_{DS}(\theta_1 \cup \theta_2) &= \frac{m_{12}(\theta_1 \cup \theta_2)}{1 - m_{12}(\emptyset)} = 0.02/0.28 = 0.0714 \end{aligned}$$

• *Advantages of DS approach*: The main interest of this DSCR-based methods lies in the fact that DSCR can be interpreted as a generalization of Bayesian conditioning and that the conditioning and the DS fusion commute, so that the three methods FC, CF or GC based all on DSCR coincide.

• *Drawbacks of DS approach*: Although attractive, DSCR approach cannot however circumvent the problem inherent to DS rule itself when the sources to combine are highly conflicting or are in worst case in total conflict. Even if the sources are not too conflicting, DSCR can yield to questionable results as pointed out in Example 1 (i.e. symmetrical results based on asymmetrical inputs) – see Table I.

C. Example 2

This example is an extension of Zadeh's example including non-existential constraint. Let's take $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ satisfying Shafer's model and the following prior bba's given in Table II, and let's assume at time $k+1$ that we learn $\theta_4 = \emptyset$, so that $\Theta_{k+1} = \{\theta_1, \theta_2, \theta_3\}$. Applying all previous methods, provide same counter-intuitive result $m_{DS}(\theta_3) = 1$ as in classical Zadeh's example.

bba's\focal elem.	θ_1	θ_2	θ_3	θ_4	$\theta_1 \cup \theta_2 \cup \theta_3$
Prior: $m_{1,k}(.)$	0.98	0	0.01	0.01	0
Prior: $m_{2,k}(.)$	0	0.98	0.01	0.01	0
Conditioning: $m_{c,k}(.)$	0	0	0	0	1
DS-FC: $m_{DS\text{-}FC}(.)$	0	0	1	0	0
DS-CF: $m_{DS\text{-}CF}(.)$	0	0	1	0	0
DS-GC: $m_{DS\text{-}GC}(.)$	0	0	1	0	0

Table II
EXAMPLE 2–A: RESULTS WITH DS-BASED CONDITIONING.

This example can be generalized as in Table III where all bba's are normalized and the non-existential constraint is $A_4 \cup \ldots \cup A_n = \emptyset$. The result of DSCR approach is given in the right column of Table III.
for $n \geq 1$, where ϵ_1, ϵ_2, and δ_{ij} are very tiny positive numbers in [0,1], a_1 and a_2 are positive numbers closer to 1, but smaller than 1, and the sum on each column is 1; all intersections $A_i \cap A_j$ are empty, where A_i can be singletons or unions of singletons. So, this is a Bayesian and non-Bayesian example.

D. Example 3

Here we give two very simple classes of examples with Bayesian or non-Bayesian bba's where DSCR cannot be applied to solve the problem. We assume Shafer's model for

Focal elem.\ bba's	$m_{1,k}(.)$	$m_{2,k}(.)$	$m_{c,k}(.)$	$m_{DS}(.)$
A_1	a_1	0	0	0
A_2	0	a_2	0	0
A_3	ϵ_1	ϵ_2	0	1
A_4	δ_{11}	δ_{21}	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_n	δ_{1n}	δ_{2n}	0	0
$A_1 \cup A_2 \cup A_3$	0	0	1	0

Table III
GENERALIZATION OF EXAMPLE 2–A.

the frames. In example 3–A, the non-existential constraint is $\theta_1 = \emptyset$ and the parameters a and b belong to $[0, 1]$.

bba's\focal elem.	θ_1	θ_2	θ_3	$\theta_2 \cup \theta_3$
Prior: $m_{1,k}(.)$	a	0	1-a	0
Prior: $m_{2,k}(.)$	b	1-b	0	0
Conditioning: $m_{c,k}(.)$	0	0	0	1

Table IV
BBA'S FOR EXAMPLE 3–A. (BAYESIAN CASE WITH $\theta_1 = \emptyset$)

Example 3–A gives 0/0 when using Dempster-Shafer's conditioning rule.

In example 3–B, we consider non-Bayesian bba's. The parameters a and b belong to $[0; 1]$. The non-existential constraint is $\theta_1 = \theta_2 = \emptyset$.

bba's\focal elem.	$\theta_1 \cup \theta_2$	θ_3	θ_4
Prior: $m_{1,k}(.)$	a	0	1-a
Prior: $m_{2,k}(.)$	b	1-b	0
Conditioning: $m_{c,k}(.)$	0	0	1

Table V
BBA'S FOR EXAMPLE 3–B. (NON-BAYESIAN CASE WITH $\theta_1 \cup \theta_2 = \emptyset$)

An infinity of Bayesian or Non Bayesian classes with total conflicting sources can be constructed where DSCR rule cannot be applied.

E. DSmT approach

Since the PCR5 or PCR6 circumvent the problem of DS rule for combining potentially highly conflicting sources of evidence, it is natural to try at first to use the same methodology for solving the problem just in replacing the DS fusion operator $\oplus$ by PCR5 (or PCR6) fusion operators. This is called PCR5CR (PCR5-based conditioning rule) or PCR6CR if one prefers to use PCR6. Unfortunately, the solution based on these PCR rules is not so simple because PCR rules are not associative and thus the result one gets highly depends on the conditioning method we adopt: FC, CF or Global. Moreover, the direct approach based on classical/original PCR5 rule under non-existential constraint cannot be applied as it will be shown from Example 1. That's why we propose a new solution to solve this important problem in the sequel.

Example 1 (continued): Let's take back example 1 and examine the results given by PCR5-FC, PCR5-CF and PCR5-GC methods[5]. The results are given in Table VI.

bba's\focal elem.	θ_1	θ_2	θ_3	$\theta_1 \cup \theta_2$
Prior: $m_{1,k}(.)$	0.2	0.4	0.3	0.1
Prior: $m_{2,k}(.)$	0.3	0.1	0.4	0.2
Conditioning: $m_{c,k}(.)$	0	0	0	1
PCR5-FC: $m_{PCR5\text{-}FC}(.)$	0.2664	0.2927	0.3320	0.1089
PCR5-CF: $m_{PCR5\text{-}CF}(.)$	0.3526	0.3822	0.0470	0.2182
PCR5-GC: $m_{PCR5\text{-}GC}(.)$	0.1811	0.1975	0.1597	0.4617

Table VI
EXAMPLE 1: RESULTS WITH PCR5-BASED CONDITIONING.

From Table VI, one sees clearly that the original PCR5 rule used for solving this example generates different results depending the method (PCR5-FC, PCR5-CF or PCR5-GC) which is not very satisfactory, and that all methods commit a positive mass to $\theta_3 = \emptyset$ which is not acceptable since we assume to work within Shafer's model in this example.

Direct approach: If we now use a direct PCR5-based approach for trying to solve the problem, we need to replace θ_3 by $\emptyset$ in the bba's inputs and apply the PCR5$_\emptyset$ fusion rule proposed in [5]. $PCR5_\emptyset$ fusion formula is same as $PCR5$ fusion formula (3) except that $X \in D^\Theta$ where D^Θ includes the empty set as well. In clear, PCR5$_\emptyset$ fusion rule allows $\emptyset$ as focal element (as in Smets' TBM). If we apply this PCR5$_\emptyset$ direct fusion, one will get results in Table VII consistent with the result of the last row of Table VI which is normal.

bba's\focal elem.	θ_1	θ_2	$\emptyset$	$\theta_1 \cup \theta_2$
Prior: $m_{1,k}(.)$	0.2	0.4	0.3	0.1
Prior: $m_{2,k}(.)$	0.3	0.1	0.4	0.2
Conditioning: $m_{c,k}(.)$	0	0	0	1
$m_{PCR5_\emptyset\text{-}Direct}(.)$	0.1811	0.1975	0.1597	0.4617

Table VII
BBA'S FOR EXAMPLE 1 AND PCR5$_\emptyset$-DIRECT RESULTS.

Example 2 (continued): Let's take back example 2 and examine the results given by PCR5-FC, PCR5-CF and PCR5-GC methods. The results are given in Table VIII (rounded when possible at the fourth decimal).

bba's\focal elem.	θ_1	θ_2	θ_3	$\theta_4 \equiv \emptyset$	$\theta_1 \cup \theta_2 \cup \theta_3$
Prior: $m_{1,k}(.)$	0.98	0	0.01	0.01	0
Prior: $m_{2,k}(.)$	0	0.98	0.01	0.01	0
Conditioning: $m_{c,k}(.)$	0	0	0	0	1
$m_{PCR5\text{-}FC}(.)$	0.49960202	0.49960202	0.00039798	0.00000016	0.00039782
$m_{PCR5\text{-}CF}(.)$	0.49970100	0.49970100	0.00049796	0.00000007	0.00009997
$m_{PCR5\text{-}GC}(.)$	0.32762253	0.32762253	0.00020045	0.00010047	0.34445402

Table VIII
EXAMPLE 2–A: RESULTS WITH PCR5-FC, PCR5-CF & PCR5-GC.

Example 3 (continued): Let's take back example 3–A with $a = b = 0.9$ and $1 - a = 1 - b = 0.1$. The results given by PCR5-FC, PCR5-CF and PCR5-GC are given in Table IX.

[5]i.e. FC, CF and GC approaches when using PCR5 rule of combination instead of DS rule.

bba's\focal elem.	$\theta_1 \equiv \emptyset$	θ_2	θ_3	$\theta_2 \cup \theta_3$
Prior: $m_{1,k}(.)$	0.9	0	0.1	0
Prior: $m_{2,k}(.)$	0.9	0.1	0	0
Conditioning: $m_{c,k}(.)$	0	0	0	1
$m_{PCR5\text{-}FC}(.)$	0.4791	0.0140	0.0140	0.4929
$m_{PCR5\text{-}CF}(.)$	0.4421	0.0605	0.0605	0.4369
$m_{PCR5\text{-}GC}(.)$	0.4435	0.0053	0.0053	0.5458

Table IX
EXAMPLE 3–A: RESULTS WITH PCR5-FC, PCR5-CF & PCR5-GC.

In summary, one has shown from very simple examples that original PCR5-based approaches cannot be used directly to solve the problem because they generate a non normalized bba (i.e. a bba with a positive value committed to $\emptyset$) and moreover the result depends the choice of the methods because of non associativity of PCR5 (or PCR6 as well). It is worth to note however that the results provided by the PCR5-based approaches commit different masses on non-empty focal lements contrariwise to DS-based approaches. In the next section we present new approaches for trying to solve the problem.

IV. EXTENDED PCR RULES

In this section we propose several ways to deal with the fusion of sources under non-existential integrity constraints since original PCR5 (or PCR6) cannot be applied directly. This is the main reason why new solutions have to be found and this is the main contribution of this paper.

A. Simple solution based on normalization

A simple solution would consist to use original PCR5 or direct PCR5$_\emptyset$ rules with a normalization final step (not included in original formulas) consisting in dividing all the mass of non-empty focal elements by $(1-m(\emptyset))$. This method can be applied only when $m(\emptyset) < 1$ of course. In example 1, one will get results given in Table X.

bba's \ focal elem.	θ_1	θ_2	θ_3	$\theta_1 \cup \theta_2$
Prior: $m_{1,k}(.)$	0.2	0.4	0.3	0.1
Prior: $m_{2,k}(.)$	0.3	0.1	0.4	0.2
Conditioning: $m_{c,k}(.)$	0	0	0	1
Normalized PCR5-FC bba	0.3988	0.4382	0	0.1630
Normalized PCR5-CF bba	0.3700	0.4010	0	0.2290
Normalized PCR5-GC bba	0.2155	0.2350	0	0.5495
Normalized PCR5$_\emptyset$ bba	0.2155	0.2350	0	0.5495
Normalized PCR6-GC bba	0.2133	0.2326	0	0.5541
Normalized PCR6$_\emptyset$ bba	0.2133	0.2326	0	0.5541

Table X
BBA'S FOR EXAMPLE 1 AND PCR5CR-BASED RESULTS AFTER NORMALIZATION.

Note that another result can be obtained from PCR5 and CF approach if one first normalizes the bba's $m_1^{PCR5}(.|\theta_1 \cup \theta_2) = m_{1,k} \oplus m_{c,k}(.)$, and $m_2^{PCR5}(.|\theta_1 \cup \theta_2) = m_{2,k} \oplus m_{c,k}(.)$ and then if we apply original PCR5 formula to combine them. We denote this method as PCR5-CnF (n standing for the position where the normalization step is done). In this case, one will get: $m_{PCR5\text{-}CnF}(\theta_1) = 0.391$, $m_{PCR5\text{-}CnF}(\theta_2) = 0.414$ and $m_{PCR5\text{-}CnF}(\theta_1 \cup \theta_2) = 0.195$ which is still different from previous results.

As one sees, all methods including a normalization step provide now different results and all agree that θ_2 corresponds to the hypothesis that has highest belief or plausibility. There is no ambiguity in the choice between θ_1 and θ_2 contrariwise to DS approach. The least uncertainty level is obtained with $PCR5-FCn$ approach in this example.

B. A more efficient solution

Here we propose another way to solve the problem using new extended PCR5 fusion formulas denoted $PCR5a$, $PCR5b$ and $PCR5c$.

• The PCR5a fusion rule: $m_{PCR5a}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \emptyset$

$$m_{PCR5a}(A) = m_{12}(A) + \sum_{\substack{X \in G^\Theta \setminus \emptyset \\ X \cap A = \emptyset}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A) + m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A) + m_1(X)}\right] + \sum_{X \in \emptyset} [m_1(A)m_2(X) + m_2(A)m_1(X)] + m_{12}(A) \cdot \frac{\sum_{X,Y \in \emptyset} m_1(X)m_2(Y)}{\sum_{Z \in G^\Theta \setminus \emptyset} m_{12}(Z)} \quad (8)$$

In PCR5a rule, one transfers the remaining conflicting masses proportionally with respect to the non-null masses resulted from the conjunctive rule.

• The PCR5b fusion rule: $m_{PCR5b}(\emptyset) = 0$ and$\forall A \in G^\Theta \setminus \emptyset$

$$m_{PCR5b}(A) = m_{12}(A) + \sum_{\substack{X \in G^\Theta \setminus \emptyset \\ X \cap A = \emptyset}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A) + m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A) + m_1(X)}\right] + \sum_{X \in \emptyset} [m_1(A)m_2(X) + m_2(A)m_1(X)] + \frac{\sum_{X,Y \in \emptyset} m_1(X)m_2(Y)}{Card(\{Z | Z \in G^\Theta \setminus \emptyset, m_{12}(Z) \neq 0\})} \quad (9)$$

In PCR5b rule, one uniformly transfers the remaining conflicting masses to all non-null masses resulted from the conjunctive rule.

• The PCR5c fusion rule: $m_{PCR5c}(\emptyset) = 0$ and$\forall A \in G^\Theta \setminus \emptyset$

$$m_{PCR5c}(A) = m_{12}(A) + \sum_{\substack{X \in G^\Theta \setminus \emptyset \\ X \cap A = \emptyset}} \left[\frac{m_1(A)^2 m_2(X)}{m_1(A) + m_2(X)} + \frac{m_2(A)^2 m_1(X)}{m_2(A) + m_1(X)}\right] + \sum_{X \in \emptyset} [m_1(A)m_2(X) + m_2(A)m_1(X)] + \sum_{X,Y \in \emptyset, A = I_t} m_1(X)m_2(Y) \quad (10)$$

In PCR5c rule, one transfers all remaining conflicting masses to the total ignorance I_t.

For PCR5a–PCR5c formulas (8)–(10) (and the next DSmHa–DSmHc, DSa–DSc formulas too) if a denominator is equal to zero, then its respective fraction is discarded, and $\sum_{X,Y\in\emptyset} m_1(X)m_2(Y)$ is transferred to the total ignorance. In this case all three PCR5a–c coincide. Similarly, all DSmHa–c coincide, and all DSa–c coincide as well. In the above formulas, $m_{12}(A)$ is the mass obtained by the classical conjunctive consensus obtained by

$$m_{12}(A) = \sum_{\substack{X_1,X_2\in G^\Theta \\ X_1\cap X_2=A}} m_1(X_1)m_2(X_2) \qquad (11)$$

G^Θ is the fusion space (power-set, hyper-power set or super-power set) depending on the underlying model chosen for the frame Θ and $\emptyset$ is the set of all empty sets that occur in the fusion due to the integrity constraints.

Remarks:

1) If no constraint occurs (i.e. no focal element becoming empty), then all PCR5a–PCR5c formulas coincide with classical PCR5 fusion rule. All these extended PCR5 rules can be extended for combining $N > 2$ sources of evidences.
2) If all information about $m_1(.)$ and $m_2(.)$ and constraints (the sets which become empty in the fusion space) come simultaneously, we can use any of these three formulas.
3) PCR5a formula is the best. PCR5a and PCR5b formulas keep the specificity resulted after applying the conjunctive rule. PCR5c rule is less specific (and not recommended).
4) These formulas can be modified easily into PCR6a–PCR6c formulas by applying PCR6 redistribution principle to $m_1(.)$ and $m_2(.)$ and transferring the remaining mass committed to empty set as in PCR5a–PCR5c formulas.
5) In the case when the information comes sequentially, we combine it in that order.

PCR5a is better than PCR5b and PCR5c because PCR5a is more specific than both of them. Its bigger specificity is due to the fact that all masses of degenerated intersections $m_{12}(A \cap B)$, where $A = B = \emptyset$, are redistributed proportionally to all non-empty elements resulted from the conjunctive rule. While PCR5c redistributes this whole degenerated mass to the total ignorance (hence the lowest specificity among this group of three related formulas), and PCR5b uniformly splits this whole degenerated mass to all non-empty elements (but this means that PCR5b gives the same amount to each non-empty element, while PCRa gives more generated mass to the elements which have a bigger mass from the conjunctive rule).

Except Smets' fusion rule in TBM, we can adapt many fusion rules which are based on the conjunctive rule, including PCR6 too of course. We can adapt in three ways, corresponding to the previous PCR5a–PCR5c improved rules, replacing only the PCR5 first summation in all three formulas with DSmH summation S_2 [4], Vol.1. For example, the DSmHa, DSmHa and DSmHc extended rules are given by:

- DSmHa fusion rule: $m_{DSmHa}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \emptyset$

$$m_{DSmHa}(A) = m_{12}(A) + \sum_{\substack{X\in G^\Theta\setminus\emptyset \\ X\cap Y=\emptyset \\ X\cup Y=A}} m_1(X)m_2(Y) + \sum_{X\in\emptyset}[m_1(A)m_2(X) + m_2(A)m_1(X)] + m_{12}(A)\cdot\frac{\sum_{X,Y\in\emptyset} m_1(X)m_2(Y)}{\sum_{Z\in G^\Theta\setminus\emptyset} m_{12}(Z)} \qquad (12)$$

- DSmHb fusion rule: $m_{DSmHb}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \emptyset$

$$m_{DSmHb}(A) = m_{12}(A) + \sum_{\substack{X\in G^\Theta\setminus\emptyset \\ X\cap Y=\emptyset \\ X\cup Y=A}} m_1(X)m_2(Y) + \sum_{X\in\emptyset}[m_1(A)m_2(X) + m_2(A)m_1(X)] + \frac{\sum_{X,Y\in\emptyset} m_1(X)m_2(Y)}{Card(\{Z|Z\in G^\Theta\setminus\emptyset, m_{12}(Z)\neq 0\})} \qquad (13)$$

- DSmHc fusion rule: $m_{DSmHc}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \emptyset$

$$m_{DSmHc}(A) = m_{12}(A) + \sum_{\substack{X\in G^\Theta\setminus\emptyset \\ X\cap Y=\emptyset \\ X\cup Y=A}} m_1(X)m_2(Y) + \sum_{X\in\emptyset}[m_1(A)m_2(X) + m_2(A)m_1(X)] + \sum_{X,Y\in\emptyset, A=I_t} m_1(X)m_2(Y) \qquad (14)$$

DSmH classic rule [4] (Vol.1) redistributes the whole conflicting mass of the form $m_{12}(A \cap B)$, with $A = B = \emptyset$, resulted from the conjunctive rule, to the total ignorance; DSmH classic is equivalent (gives the same result) as DSmHc. But DSmHa and DSmHb are more specific than DSmHc (=DSmH classic) from exactly the same reason as explained before regarding the more specificity of PCR5a with respect to PCR5 and PCR5b. DSmHa is the most specific among all three DSmHa–DSmHc. That's why we need DSmHa. In addition, in the three formulas of DSmHa–DSmHc we can condensed the first two summations $(m_{12}(A) + \sum \ldots + \sum \ldots + \ldots)$ into one summation only, i.e under the first summation we can write $X, Y \in G^\Theta$ (so X, Y can be empty as well) and the second summation disappears (it is absorbed by the first).

Similarly for Dempster-Shafer's extended rule in the DSm way, we replace in all first three formulas the first PCR5 summation by

$$m_{12}(A)\cdot\frac{\sum_{\substack{X,Y\in\emptyset \\ X\cap Y=\emptyset}} m_1(X)m_2(Y)}{\sum_{Z\in G^\Theta\setminus\emptyset} m_{12}(Z)}$$

• DSa fusion rule: $m_{DSa}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \emptyset$

$$m_{DSa}(A) = m_{12}(A) + m_{12}(A) \cdot \frac{\sum_{\substack{X,Y\in\emptyset \\ X\cap Y=\emptyset}} m_1(X)m_2(Y)}{\sum_{Z\in G^\Theta\setminus\emptyset} m_{12}(Z)} + \sum_{X\in\emptyset}[m_1(A)m_2(X) + m_2(A)m_1(X)] + m_{12}(A) \cdot \frac{\sum_{X,Y\in\emptyset} m_1(X)m_2(Y)}{\sum_{Z\in G^\Theta\setminus\emptyset} m_{12}(Z)} \quad (15)$$

• DSb fusion rule: $m_{DSb}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \emptyset$

$$m_{DSb}(A) = m_{12}(A) + m_{12}(A) \cdot \frac{\sum_{\substack{X,Y\in\emptyset \\ X\cap Y=\emptyset}} m_1(X)m_2(Y)}{\sum_{Z\in G^\Theta\setminus\emptyset} m_{12}(Z)} + \sum_{X\in\emptyset}[m_1(A)m_2(X) + m_2(A)m_1(X)] + \frac{\sum_{X,Y\in\emptyset} m_1(X)m_2(Y)}{Card(\{Z|Z \in G^\Theta \setminus \emptyset, m_{12}(Z) \neq 0\})} \quad (16)$$

• DSc fusion rule: $m_{DSc}(\emptyset) = 0$ and $\forall A \in G^\Theta \setminus \emptyset$

$$m_{DSc}(A) = m_{12}(A) + m_{12}(A) \cdot \frac{\sum_{\substack{X,Y\in\emptyset \\ X\cap Y=\emptyset}} m_1(X)m_2(Y)}{\sum_{Z\in G^\Theta\setminus\emptyset} m_{12}(Z)} + \sum_{X\in\emptyset}[m_1(A)m_2(X) + m_2(A)m_1(X)] + \sum_{X,Y\in\emptyset, A=I_t} m_1(X)m_2(Y) \quad (17)$$

Note that all these extended fusion rules are however not associative and therefore if one has several sources available at a given time to combine, the combination must be applied with all sources together to get optimal fusion result.

C. Example 1 (continued)

Let's examine in details the results obtained on Example 1 with all these extended fusion formulas. Because $\theta_3 = \emptyset$ and Shafer's model is assumed for Θ_k, the set of elements becoming empty is $\emptyset = \{\theta_1 \cap \theta_2, \theta_1 \cap \theta_3, \theta_2 \cap \theta_3, (\theta_1 \cup \theta_2) \cap \theta_3, \theta_3\}$ and one has: $m_{12}(\theta_1 \cap \theta_2) = 0.14$, $m_{12}(\theta_1 \cap \theta_3) = 0.17$, $m_{12}(\theta_2 \cap \theta_3) = 0.19$, $m_{12}((\theta_1 \cup \theta_2) \cap \theta_3) = 0.10$, $m_{12}(\theta_3) = 0.12$. $m_{12}(\theta_1 \cap \theta_2 \in \emptyset) = 0.14$ is redistributed back to θ_1 and θ_2 using PCR5 principle:

$$\frac{x_{1\theta_1}}{0.2} = \frac{y_{1\theta_2}}{0.1} = \frac{0.02}{0.3} = \frac{0.2}{3}, \qquad \frac{x_{2\theta_1}}{0.3} = \frac{y_{2\theta_2}}{0.4} = \frac{0.12}{0.7} = \frac{1.2}{7}$$

$$x_{1\theta_1} = 0.2\frac{0.2}{3} \approx 0.013 \qquad x_{2\theta_1} = 0.3\frac{1.2}{7} \approx 0.051$$

$$y_{1\theta_2} = 0.1\frac{0.2}{3} \approx 0.007 \qquad y_{2\theta_2} = 0.4\frac{1.2}{7} \approx 0.069$$

$m_{12}(\theta_1 \cap \theta_3 \in \emptyset) = 0.17$ is all redistributed back to θ_1 since $\theta_3 = \emptyset$ (non-existential constraint). $m_{12}(\theta_2 \cap \theta_3 \in \emptyset) = 0.19$ is all redistributed back to θ_2 since $\theta_3 = \emptyset$ (non-existential constraint). $m_{12}((\theta_1 \cup \theta_2) \cap \theta_3) \in \emptyset) = 0.10$ is all redistributed back to $\theta_1 \cup \theta_2$ since $\theta_3 = \emptyset$ (non-existential constraint). While $m_{12}(\theta_3 = \emptyset) = 0.12$ is redistributed differently in each PCR5a, PCR5b and PCR5c formulas:

1) In PCR5a:

$$\frac{x_{\theta_1}}{0.13} = \frac{y_{\theta_2}}{0.13} = \frac{z_{\theta_1\cup\theta_2}}{0.02} = \frac{0.12}{0.28} = \frac{3}{7}$$

whence $x_{\theta_1} = y_{\theta_2} = 0.13 \cdot 3/7 \approx 0.056$ and $z_{\theta_1\cup\theta_2} = 0.02 \cdot 3/7 \approx 0.008$.

2) In PCR5b: $x_{\theta_1} = y_{\theta_2} = z_{\theta_1\cup\theta_2} = 0.12/3 = 0.04$.
3) In PCR5c: $z_{\theta_1\cup\theta_2} = 0.12$.

Finally, one then gets results shown in the Table XI. From these results, one sees that PCR5a rules provides the most specific result since the mass committed to the uncertainty is lowest with respect to what we get with PCR5b, PCR5c and other PCR5-based normalized conditioning rules given in the Table X. PCR5b is also a bit better (more specific) than PCR5-based normalized conditioning rules also. As we see and as expected from the theory PCR5c is less specific than PCR5a and PCR5b. If we use DSmHa-DSmHc fusion rules on this example, $m_{12}(\theta_1 \cap \theta_2 \in \emptyset) = 0.14$ is all redistributed back to $\theta_1 \cup \theta_2$ using DSmH principle [4], Vol.1. The other conflicting masses are redistributed respectively in the same way in PCR5a–PCR5c rules. The same example for Dempster-Shafer's rule extended in DSm style: $m_{12}(\theta_1 \cap \theta_2 \in \emptyset) = 0.14$ is all redistributed back to θ_1, θ_2, and $\theta_1 \cup \theta_2$ since they are non-empty proportionally with respect to their conjunctive rule masses 0.13, 0.13 and respectively 0.02:

$$\frac{x_{\theta_1}}{0.13} = \frac{y_{\theta_2}}{0.13} = \frac{z_{\theta_1\cup\theta_2}}{0.02} = \frac{0.14}{0.28} = 0.5$$

whence $x_{\theta_1} = y_{\theta_2} = 0.13(0.5) = 0.065$ and $z_{\theta_1\cup\theta_2} = 0.02(0.5) = 0.010$. The other conflicting masses are redistributed respectively in the same way as in PCR5a–PCR5c rules. The results obtained with DSmHa–DSmHc and DSa–DSc rules are given in Table XI. In this example, one sees that PCR5a is the most specific rule and in all cases, the rational decision to take will be θ_2 without ambiguity contrariwise to DSCR approach.

bba's \ focal elem.	θ_1	θ_2	$\theta_3 \equiv \emptyset$	$\theta_1 \cup \theta_2$
Prior: $m_{1,k}(.)$	0.2	0.4	0.3	0.1
Prior: $m_{2,k}(.)$	0.3	0.1	0.4	0.2
m_{PCR5a}	**0.420**	**0.452**	0	**0.128**
m_{PCR5b}	0.404	0.436	0	0.160
m_{PCR5c}	0.364	0.396	0	0.240
m_{DSa}	0.421	0.441	0	0.138
m_{DSb}	0.405	0.425	0	0.170
m_{DSc}	0.365	0.385	0	0.250
m_{DSmHa}	0.356	0.376	0	0.268
m_{DSmHb}	0.340	0.360	0	0.300
m_{DSmHc}	0.300	0.320	0	0.380

Table XI
EXAMPLE 1: PCR5A–C & DSA–C & DSMHA–C RESULTS.

V. EXAMPLES

Here we present the solution of Examples 2–A, 3A–3B obtained with our new extended PCR5a–PCR5c rules of combination for solving the fusion of bba's under non-existential constraints in degenerate cases.

A. Example 2 (continued)

Let's consider the Example 2–A and apply PCR5a–PCR5c formulas. Using the PCR5 principle, $m_{12}(\theta_1 \cap \theta_2) = 0.98 \cdot 0.98 = 0.9604$ is redistributed back to θ_1 and θ_2 with the same proportions $x_{\theta_1} = x_{\theta_2} = 0.4802$; $m_{12}(\theta_1 \cap \theta_3) = 0.98 \cdot 0.01 = 0.0098$ is redistributed to θ_1 and θ_3 with $x_{\theta_1} = 0.00970101$ and $x_{\theta_3} = 0.00009899$; $m_{12}(\theta_2 \cap \theta_3) = 0.0098$ is redistributed to θ_2 and θ_3 with $x_{\theta_2} = 0.00970101$ and $x_{\theta_3} = 0.00009899$; $m_{12}(\theta_1 \cap \theta_4) = 0.0098$ is transferred to θ_1 only since $\theta_4 \equiv \emptyset$; $m_{12}(\theta_2 \cap \theta_4) = 0.0098$ is transferred to θ_2 only since $\theta_4 \equiv \emptyset$; $m_{12}(\theta_3 \cap \theta_4) = 0.0002$ is transferred to θ_3 only since $\theta_4 \equiv \emptyset$; Since only $m_{12}(\theta_3) \neq 0$ with $\theta_3 \neq \emptyset$ the mass $m_{12}(\theta_4) = 0.0001$ is transferred to θ_3 in both PCR5a and PCR5b formulas. But in PCR5c rule, $m_{12}(\theta_4)$ is transferred to the total ignorance $I_t = \theta_1 \cup \theta_2 \cup \theta_3$. The final results obtained with PCR5a, PCR5b (same as with PCR5a for this example) and PCR5c are given in Table XII below.

focal el.\bba's	$m_{1,k}$	$m_{2,k}$	$m_{PCR5a,b}(.)$	$m_{PCR5c}(.)$
θ_1	0.98	0	0.49970101	0.49970101
θ_2	0	0.98	0.49970101	0.49970101
θ_3	0.01	0.01	$5.9798 \cdot 10^{-4}$	$4.9798 \cdot 10^{-4}$
$\theta_4 \equiv \emptyset$	0.01	0.01	0	0
$\theta_1 \cup \theta_2 \cup \theta_3$	0	0	0	0.0001

Table XII
EXAMPLE 2–A: RESULTS WITH PCR5A–C

B. Example 3 (continued)

In Example 3–A, θ_1 becomes empty and therefore: $m_{12}(\theta_1 \cap \theta_2) = a(1-b)$ goes to θ_2, $m_{12}(\theta_1 \cap \theta_3) = b(1-c)$ goes to θ_3 and $m_{12}(\theta_2 \cap \theta_3) = 1 - a - b + ab$ is split between θ_2 and θ_3 proportionally to $1-b$ and $1-a$ respectively:

$$\frac{x_{\theta_2}}{1-b} = \frac{x_{\theta_3}}{1-a} = \frac{1-a-b+ab}{2-a-b}$$

Therefore, one gets finally

$$x_{\theta_2} = \frac{1-a-2b+ab+b^2-ab^2}{2-a-b}$$

$$x_{\theta_3} = \frac{1-2a-b+2ab+a^2-a^2b}{2-a-b}$$

Since $\theta_1 = \emptyset$, $m_{12}(\theta_1) = ab$ is redistributed to $\theta_2 \cup \theta_3$ in PCR5a–PCR5c formulas because all $m_{12}(X) = 0$ for $X \neq \emptyset$. The final results are given in Table XIII depending on the values of parameters a and b

Cases	$a \neq 1, b \neq 1$	$a = b = 1$
focal elem. \ bba's	$m_{PCR5a,b,c}(.)$	$m_{PCR5a,b,c}(.)$
θ_1	0	0
θ_2	$a(1-b) + \frac{(1-b)(1-a-b+ab}{2-a-b}$	0
θ_3	$b(1-a) + \frac{(1-a)(1-a-b+ab}{2-a-b}$	0
$\theta_2 \cup \theta_3$	ab	1

Table XIII
EXAMPLE 3–A: RESULTS WITH PCR5A–PCR5C

Extended PCR5 rules for Example 3–B give same results as for Example 3–A, where we replace θ_1 by $\theta_1 \cup \theta_2$, θ_2 by θ_3, and θ_3 by θ_4, and $\theta_2 \cup \theta_3$ by $\theta_3 \cup \theta_4$. If we take by example, $a = b = 0.9$ and $1 - a = 1 - b = 0.1$ in examples 3–A and 3–B then we will finally obtain for Examples 3–A & 3–B:

bba's\focal elem.	θ_1	θ_2	θ_3	$\theta_2 \cup \theta_3$
$m_{PCR5a-c}(.)$	0	0.095	0.095	0.810
$m_{DSmHa-c}(.)$	0	0.090	0.090	0.820
$m_{DSa-c}(.)$	0	0.090	0.090	0.820

Table XIV
EXAMPLE 3–A: RESULTS WITH $a = b = 0.9$ AND $1 - a = 1 - b = 0.1$

bba's\focal elem.	$\theta_1 \cup \theta_2$	θ_3	θ_4	$\theta_3 \cup \theta_4$
$m_{PCR5a-c}(.)$	0	0.095	0.095	0.810
$m_{DSmHa-c}(.)$	0	0.090	0.090	0.820
$m_{DSa-c}(.)$	0	0.090	0.090	0.820

Table XV
EXAMPLE 3–B: RESULTS WITH $a = b = 0.9$ AND $1 - a = 1 - b = 0.1$

Dempster-Shafer's rule cannot be applied in these examples since it gives 0/0.

VI. CONCLUSIONS

In this paper we extend the classical PCR5 and DSmH combination fusion rules to two ensembles of new fusion rule formulas, PCR5a–PCR5c and respectively DSmHa–DSmHc, in order to be able to take into consideration the non-existence constraints (i.e. when some sets become empty) that may occur during a dynamic fusion. Further, we show that the same DSmT extension procedure applied to PCR5 and DSmH can be applied to Dempster's rule and other rules as well. We provide several examples with these PCR5a–PCR5c and DSmHa–DSmHc rules, and also with Dempster-Shafer conditioning rule (DSCR). We have presented some classes of counter-examples to DSCR. If we have two sources, what to do first Fusion and then Conditioning, or Conditioning and then Fusion? A simple answer would be to do them in the order we receive the information. But in the case we receive all of them simultaneously, it is better to use these new extended rules depending on the specificity quality we want to get, PCR5a being the most specific rule.

REFERENCES

[1] J. Dezert, F. Smarandache, *A new probabilistic transformation of belief mas assignment*, in Proc. of Fusion 2008, Cologne, Germany, July 2008. (available from http://fs.gallup.unm.edu//DSmT.htm).

[2] J. Dezert, F. Smarandache, *Non Bayesian conditioning and deconditioning*, International Workshop on Belief Functions, Brest, France, April 2010. http://bfas.iutlan.univ-rennes1.fr/belief2010/

[3] G.Shafer, *A mathematical theory of evidence*, Princeton Univ. Press, 1976.

[4] F. Smarandache, J. Dezert (Editors), *Advances and Applications of DSmT for Information Fusion*, American Research Press, Rehoboth, Vol.1-3, 2004-2009. (available from http://fs.gallup.unm.edu//DSmT.htm)

[5] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, in Proceedings of Fusion 2010 conference, Edinburgh, UK, July 2010.

[6] P. Smets, *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in Artificial Intelligence, Vol. 5, pp. 29–39, 1990.

[7] http://bfas.iutlan.univ-rennes1.fr/wiki/index.php/Toolboxs

A Fuzzy-Cautious OWA Approach with Evidential Reasoning

Deqiang Han
Jean Dezert
Jean-Marc Tacnet
Chongzhao Han

Originally published as Han D., Dezert J., Tacnet J.-M., Han C., *A Fuzzy-Cautious OWA Approach with Evidential Reasoning*, in Proc. Of Fusion 2012, Singapore, July 2012, and reprinted with permission.

***Abstract*—Multi-criteria decision making (MCDM) is to make decisions in the presence of multiple criteria. To make a decision in the framework of MCDM under uncertainty, a novel fuzzy - Cautious OWA with evidential reasoning (FCOWA-ER) approach is proposed in this paper. Payoff matrix and belief functions of states of nature are used to generate the expected payoffs, based on which, two Fuzzy Membership Functions (FMFs) representing optimistic and pessimistic attitude, respectively can be obtained. Two basic belief assignments (bba's) are then generated from the two FMFs. By evidence combination, a combined bba is obtained, which can be used to make the decision. There is no problem of weights selection in FCOWA-ER as in traditional OWA. When compared with other evidential reasoning-based OWA approaches such as COWA-ER, FCOWA-ER has lower computational cost and clearer physical meaning. Some experiments and related analyses are provided to justify our proposed FCOWA-ER.**

***Index Terms*—Evidence theory, OWA, belief function, uncertainty, decision making, information fusion.**

I. Introduction

In real-life situations, decision making always encounters difficult multi-criteria problems [1]. In classical Multi-Criteria Decision Making (MCDM) framework, the ordered weighted averaging (OWA) approach proposed by Yager [2] has been increasingly used in wide range of successful applications for the aggregation of decision making problems such as image processing, fuzzy control, market prediction and expert systems, etc [3]. OWA is a generalized mean operator providing flexibility in the aggregation. Thus the aggregation can be bounded between minimum and maximum operators. This flexibility of the OWA operator is implemented by using the concept of orness (optimism) [4], which is a surrogate for decision maker's attitude. One important issue in the OWA aggregation is the determination of the associated weights. Many approaches [5]–[10] have been proposed to determine the weights in OWA. See the related references for details.

In multi-criteria decision making, decisions are often made under uncertainty, which are provided by several more or less reliable sources and depend on the states of the world: decisions can be taken in certain, risky or uncertain environment. To implement the decision making under uncertainty, many approaches were proposed including DS-AHP [11], DSmT-AHP [12] and ER-MCDA [13], etc. Especially for the OWA under uncertainty, Yager proposed an OWA approach with evidence reasoning [14]. In our previous work, a cautious OWA with evidential reasoning (COWA-ER) was proposed to take into account the imperfect evaluations of the alternatives and the unknown beliefs about groups of the possible states of the world. COWA-ER mixes MCDM principles, decision under uncertainty principles and evidential reasoning. There is no step of weights selection in COWA-ER, which is good for the practical use. Recently, we find that there also exists drawbacks in COWA-ER. More precisely, the computational cost of the combination of different evidences by COWA- ER highly depends on the number of alternatives we encounter in decision making. When the number of alternatives is large, the computational cost will increase significantly.

In this paper, we propose a modified COWA-ER approach, called Fuzzy-Cautious OWA with Evidential Reasoning (FCOWA-ER), by using a different way to manage the uncertainty caused by weights selection. Payoff matrix together with the belief structure (knowledge of the states of the nature) are used to generate two Fuzzy Membership Functions (FMFs) representing the optimistic and pessimistic attitude, respectively. Then two bba's can be obtained based on the two FMF's by using α-cut approach. Based on evidence combination, the combined bba can be obtained and the final decision can be made. The FCOWA-ER approach doesn't need a (ad-hoc) selection of weights as in the traditional OWA. When compared with COWA-ER, FCOWA-ER has less computational cost and clearer physical meaning because it requires only one combination operation regardless of the number of alternatives. The proposed FCOWA-ER can be seen as a trade-off between the optimistic and the pessimistic attitudes. The preference of the two attitudes can be adjusted by the users using discounting factors in the combination of evidences. Some experiments and related analyses are provided to show the rationality and efficiency of this new FCOWA-ER approach.

II. Multi-Criteria Decision Making under Uncertainty

Multi-criteria decision making (MCDM) refers to making decisions in the presence of multiple, usually conflicting or discordant, criteria. Consider the following matrix C provided to a decision maker:

$$\begin{array}{c} \\ A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{array} \begin{array}{c} \begin{array}{cccccc} S_1 & \cdots & S_j & \cdots & S_n \end{array} \\ \left[\begin{array}{ccccc} C_{11} & \cdots & C_{1j} & \cdots & C_{1n} \\ \vdots & & \vdots & & \vdots \\ C_{i1} & & C_{ij} & & C_{in} \\ \vdots & & \vdots & & \vdots \\ C_{q1} & \cdots & C_{qj} & \cdots & C_{qn} \end{array}\right] \end{array} = C$$

In the above each A_i corresponds to a possible alternative available to the decision maker. Each S_j corresponds to a possible value of the variable called the state of nature. C_{ij} corresponds to the payoff to be received by the decision maker if he selects action A_i and the state of nature is S_j. The problem encountered by the decision maker in MCDM is to select the action which gives him the optimum payoff.

Among all the available MCDM approaches, Ordered Weighted Averaging (OWA) is a very important one, which is introduced below.

A. Ordered Weighted Averaging (OWA)

OWA was proposed by Yager in [2]. An OWA operator of dimension n is a function $F : \mathbb{R}^n \to \mathbb{R}$ that has associated with a weighting vector[1] $W = [w_1, w_2, ..., w_n]^T$ such that $w_i \in [0, 1]$ and $\sum_{i=1}^{n} w_i = 1$. For any set of values $a_1, ..., a_n$

$$F(a_1, ..., a_n) = \sum\nolimits_{i=1}^{n} (w_i \cdot b_i) \tag{1}$$

where b_i is the ith largest element in the collection $a_1, ..., a_n$. It should be noted that the weights in the OWA operator are associated with a position in the ordered arguments rather than a particular argument.

The OWA operator depends on the associated weights, hence the weights determination is very crucial. Some commonly used weights selection strategies are as follows [14]:

1) Pessimistic Attitude: If $W = [0, 0, ..., 1]^T$, then

$$F(a_1, a_2, ..., a_n) = \min_j [a_j].$$

2) Optimistic Attitude: If $W = [1, 0, ..., 0]^T$, then

$$F(a_1, a_2, ..., a_n) = \max_j [a_j].$$

3) Hurwicz Strategy: If $W = [\alpha, 0, ..., 1 - \alpha]^T$, then

$$F(a_1, a_2, ..., a_n) = \alpha \cdot \max_j [a_j] + (1 - \alpha) \cdot \min[a_j].$$

4) Normative Strategy: If $W = [1/n, 1/n, ..., 1/n]^T$, then

$$F(a_1, a_2, ..., a_n) = (1/n) \cdot \sum\nolimits_{i=1}^{n} a_i.$$

The OWA operator can be seen as the decision-making under ignorance, because in classical OWA, there is no knowledge about the true state of the nature but that it belongs to a finite set. It should be noted that the pessimistic and optimistic strategies provide limited classes of OWA operators. There also exist other strategies to determine the weights, e.g., the weights generation based on entropy maximization. See related references [5]–[10] for details.

[1] where X is a vector or a matrix and X^T denotes the transpose of X.

Based on such OWA operators, for each alternative A_i, $i = 1, \ldots, q$, we can choose a weighting vector $W_i = [w_{i1}, w_{i2}, \ldots w_{in}]$ and compute its OWA value $V_i \triangleq F(C_{i1}, C_{i2}, \ldots, C_{in}) = \sum_j w_{ij} \cdot b_{ij}$ where b_{ij} is the jth largest element in the collection of payoffs $C_{i1}, C_{i2}, \ldots, C_{in}$. Then, as for decision-making under ignorance, we choose $A^* = A_{i^*}$ with $i^* \triangleq \arg\max_i \{V_i\}$.

B. Uncertainty in MCDM context

Decisions are often made based on imperfect information and knowledge (imprecise, uncertain, incomplete) provided by several more or less reliable sources and depend on the states of the world: decisions can be taken in certain, risky or uncertain environment [15]. In a MCDM context, the decision under uncertainty means that the evaluations of the alternative are dependent on the state of the world.

Introducing the ignorance and the uncertainty in a MCDM process consists in considering that consequences of alternatives (A_i) depend on the state of nature represented by a finite set $S = \{S_1, S_2, ..., S_n\}$. For each state, the MCDM method provides an evaluation C_{ij}. We assume that this evaluation C_{ij} done by the decision maker corresponds to the choice of $A_i \in \{A_1, ..., A_q\}$ when S_j occurs with a given (possibly subjective) probability. The evaluation matrix is defined as $C = [C_{ij}]$ where $i = 1, ..., q$ and $j = 1, ..., n$.

Since the payoff to the decision maker depends upon the state of nature, his procedure for selecting the best alternative depends upon the type of knowledge he has about the state of nature. For representing the uncertainty for the state of nature, the belief functions introduced in Dempster-Shafer Theory (DST) [16] (known also as the Evidence Theory) can be used. This is briefly introduced below.

C. Basics of Evidence Theory

In DST, the elements in the frame of discernment (FOD) denoted by Θ are mutually exclusive and exhaustive. Suppose 2^Θ denotes the powerset of FOD. One defines the function $m : 2^\Theta \to [0, 1]$ as the basic belief assignment (bba, also called mass function) if it satisfies:

$$\sum\nolimits_{A \subseteq \Theta} m(A) = 1, \; m(\emptyset) = 0 \tag{2}$$

The belief function (Bel) and the plausibility function (Pl) are defined below, respectively:

$$Bel(A) = \sum\nolimits_{B \subseteq A} m(B) \tag{3}$$

$$Pl(A) = \sum\nolimits_{A \cap B \neq \emptyset} m(B) \tag{4}$$

Let us consider two bba's $m_1(.)$ and $m_2(.)$ defined over the FOD Θ. Their corresponding focal elements[2] are $A_1, ..., A_k$ and $B_1, ..., B_l$. If $k = \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) < 1$, the function $m : 2^\Theta \to [0, 1]$ denoted by

$$m(A) = \begin{cases} 0, \; A = \emptyset \\ \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)}, A \neq \emptyset \end{cases} \tag{5}$$

[2] a focal element X of a bba $m(.)$ is an element of the power set of the FOD such that $m(X) > 0$.

is also a bba. The rule defined in Eq. (5) is called Dempster's rule of combination.

D. MCDM with belief structures

Yager proposed an approach for decision making with belief structures [14]. One considers a collection of q alternatives belonging to $A = \{A_1, A_2, ..., A_q\}$ and a finite set $S = \{S_1, S_2, ..., S_n\}$ of states of the nature. We assume that the payoff/gain C_{ij} of the decision maker in choosing A_i when S_j occurs are given by positive (or null) numbers. The payoffs $q \times n$ matrix is defined by $C = [C_{ij}]$ where $i = 1, ..., q$ and $j = 1, ..., n$ as in eq. (2). The decision-making problem consists in choosing the alternative $A^* \in A$ which maximizes the payoff to the decision maker given the knowledge on the state of the nature and the payoffs matrix C. $A^* \in A$ is called the best alternative or the solution (if any) of the decision-making problem.

In Yager's approach, the knowledge on the state of the nature is characterized by a belief structure. Clearly, one assumes that a priori knowledge on the frame S of the different states of the nature is given by a bba $m(.) : 2^S \to [0, 1]$. Decision under certainty is characterized by $m(S_j) = 1$; Decision under risk is characterized by $m(S_j) > 0$ for some states $S_j \in S$; Decision under full ignorance is characterized by $m(S_1 \cup S_2 \cup ... \cup S_n) = 1$, etc. Yager's OWA for decision making under uncertainty combines the schemes used for decision making under risk and ignorance. It is based on the derivation of a generalized expected value C_i of payoff for each alternative A_i as follows:

$$C_i = \sum\nolimits_{k=1}^{r} m(X_k) V_{ik} \tag{6}$$

where r is the number of focal elements of the belief structure. $m(X_k)$ is the mass assignment of the focal element $X_k \in 2^S$. V_{ik} is the payoff we get when we select alternative A_i and the state of nature lies in X_k. The derivation of V_{ik} is done similarly as for the decision making under ignorance (i.e., the procedure of OWA) when restricting the states of the nature to the subset of states belonging to X_k only. One can choose different strategies to determinate the weights. Actually, C_i is essentially the expected value of the payoffs under A_i. Select the alternative with highest C_i as the optimal one.

E. Cautious OWA with Evidential Reasoning

Yager's OWA approach is based on the choice of a given attitude measured by an optimistic index in [0, 1] to get the weighting vector W. How to choose such an index/attitude? This choice is ad-hoc and very disputable for users. In our previous work [15] we have only considered jointly the two extreme attitudes (pessimistic and optimistic ones) jointly and developed a method called Cautious OWA with Evidential Reasoning (COWA-ER) for decision under uncertainty based on the imprecise evaluation of alternatives.

In COWA-ER, the pessimistic and optimistic OWA are used respectively to construct the intervals of expected payoffs for different alternatives. For example, if there exist q alternatives, the expected payoffs are as follows.

$$E[C] = \begin{bmatrix} E[C_1] \\ E[C_2] \\ \vdots \\ E[C_q] \end{bmatrix} = \begin{bmatrix} [C_1^{\min}, C_1^{\max}] \\ [C_2^{\min}, C_2^{\max}] \\ \vdots \\ [C_q^{\min}, C_q^{\max}] \end{bmatrix}$$

Therefore, one has q sources of information about the parameter associated with the best alternative to choose. For decision making under imprecision, the belief functions framework is used again. COWA-ER includes four steps:

- Step 1: normalization of imprecise values in $[0, 1]$;
- Step 2: conversion of each normalized imprecise value into elementary bba $m_i(.)$;
- Step 3: fusion of bba $m_i(.)$ with some combination rule;
- Step 4: choice of the final decision based on the resulting combined bba.

In step 2, we convert each imprecise value into its bba according to a very natural and simple transformation [17]. Here, we need to consider the finite set of alternatives $\Theta = \{A_1, A_2, \ldots, A_q\}$ as the frame of discernment and the sources of belief associated with them are obtained from the normalized imprecise expected payoff vector $E^{Imp}[C]$. The modeling for computing a bba associated to A_i from any imprecise value $[a; b] \subseteq [0; 1]$ is simple and is done as follows:

$$\begin{cases} m_i(A_i) = a, \\ m_i(\bar{A}_i) = 1 - b \\ m_i(A_i \cup \bar{A}_i) = m_i(\Theta) = b - a \end{cases} \tag{7}$$

where $\bar{A}_i$ is the A_i's complement in Θ. With such a conversion, one sees that $Bel(A_i) = a$, $Pl(A_i) = b$. The uncertainty is represented by length of the interval $[a; b]$ and corresponds to the imprecision of the variable (here the expected payoff) on which the belief function for A_i is defined.

III. A Novel Fuzzy-COWA-ER

The COWA-ER has its rationality and can well process the MCDM under uncertainty. However the complexity and the computational time of the combination of COWA-ER method is highly dependent on the number of alternatives used for decision-making. When the number of alternatives is large, the computational cost will increase significantly. In COWA-ER, each expected interval is used as the information sources, however, these expected intervals are jointly obtained and thus these information sources are relatively correlated, which is harmful for the followed evidence combination. In this paper, we propose modified COWA-ER called Fuzzy-COWA-ER. Before presenting the principle of FCOWA-ER, we first recall that the pessimistic and optimistic OWA versions are used respectively to construct the intervals of expected payoffs for different alternatives as follows:

$$E[C] = \begin{bmatrix} E[C_1] \\ E[C_2] \\ \vdots \\ E[C_q] \end{bmatrix} = \begin{bmatrix} [C_1^{\min}, C_1^{\max}] \\ [C_2^{\min}, C_2^{\max}] \\ \vdots \\ [C_q^{\min}, C_q^{\max}] \end{bmatrix}$$

A. Principle of FCOWA-ER

In COWA-ER, each row of the expected payoff $E[C]$ is used as information sources while in FCOWA-ER, we consider the two columns of $E[C]$ as two information sources, representing the pessimistic and the optimistic attitude, respectively. The column-wise normalized expected payoff is

$$E^{Fuzzy}[C] = \begin{bmatrix} N_1^{\min}, N_1^{\max} \\ N_2^{\min}, N_2^{\max} \\ \vdots \\ N_q^{\min}, N_q^{\max} \end{bmatrix}$$

where $N_i^{\min} \in [0,1]$ $(i = 1, ..., q)$ represents the normalized value in the column of pessimistic attitude and $N_i^{\max} \in [0,1]$ represents the normalized value in the column of optimistic attitude. The vectors $[N_1^{\min}, ..., N_q^{\min}]$ and $[N_1^{\max}, ..., N_q^{\max}]$ can be seen as two fuzzy membership functions (FMFs) representing the possibilities of all the alternatives: $A_1, ..., A_q$.

The principle of FCOWA-ER includes the following steps:

- Step 1: normalize each column in $E[C]$, respectively, to obtain $E^{Fuzzy}[C]$;
- Step 2: conversion of two normalized columns, i.e., two FMFs into two bba's $m_{Pess}(.)$ and $m_{Opti}(.)$;
- Step 3: fusion of bba's $m_{Pess}(.)$ and $m_{Opti}(.)$ with some combination rule;
- Step 4: choice of the final decision based on the resulting combined bba.

In Step 2, we implement the conversion of the FMF into the bba by using α-cut approach as follows:

Suppose the FOD is $\Theta = \{A_1, A_2, ..., A_q\}$ and the FMF is $\mu(A_i), i = 1, ..., q$, the corresponding bba introduced in [18] is used to generate M α-cut ($0 < \alpha_1 < \alpha_2 < \cdots < \alpha_M \leq 1$), where $M \leq |\Theta| = n$.

$$\begin{cases} B_j = \{A_i \in \Theta | \mu(A_i) \geq \alpha_j\} \\ m(B_j) = \frac{\alpha_j - \alpha_{j-1}}{\alpha_M} \end{cases} \quad (8)$$

B_j, for $j = 1, ..., M$, $(M \leq |\Theta|)$ represents the focal element. For simplicity, here we set $M = q$ and $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_q \leq 1$ as the sort of $\mu(A_i)$.

B. Example of FCOWA-ER versus COWA-ER and OWA

Example 1: Let's take states $S = \{S_1, S_2, S_3, S_4, S_5\}$ with the associated bba $m(.)$ given by:

$$\begin{cases} m(S_1 \cup S_3 \cup S_4) = 0.6 \\ m(S_2 \cup S_5) = 0.3 \\ m(S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5) = 0.1 \end{cases}$$

Let's also consider alternatives $A = \{A_1, A_2, A_3, A_4\}$ and the payoffs matrix:

$$C = \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} \quad (9)$$

1) Implementation of OWA: The $r = 3$ focal elements of $m(.)$ are $X_1 = S_1 \cup S_3 \cup S_4$, $X_2 = S_2 \cup S_5$ and $X_3 = S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5$. X_1 and X_2 are partial ignorance and X_3 is the full ignorance. One considers the following submatrix (called bags by Yager) for the derivation of V_{ik}, for $i = 1, 2, 3, 4$ and $k = 1, 2, 3$.

$$M(X_1) = \begin{bmatrix} M_{11} \\ M_{21} \\ M_{31} \\ M_{41} \end{bmatrix} = \begin{bmatrix} 7 & 12 & 13 \\ 12 & 5 & 11 \\ 9 & 3 & 10 \\ 6 & 11 & 15 \end{bmatrix}$$

$$M(X_2) = \begin{bmatrix} M_{12} \\ M_{22} \\ M_{32} \\ M_{42} \end{bmatrix} = \begin{bmatrix} 5 & 6 \\ 10 & 2 \\ 13 & 9 \\ 9 & 4 \end{bmatrix}$$

$$M(X_3) = \begin{bmatrix} M_{13} \\ M_{23} \\ M_{33} \\ M_{43} \end{bmatrix} = \begin{bmatrix} 7 & 5 & 12 & 13 & 6 \\ 12 & 10 & 5 & 11 & 2 \\ 9 & 13 & 3 & 10 & 9 \\ 6 & 9 & 11 & 15 & 4 \end{bmatrix} = C$$

• Using pessimistic attitude, and applying the OWA operator on each row of $M(X_k)$ for $k = 1$ to r, one gets finally: $V(X_1) = [V_{11}, V_{21}, V_{31}, V_{41}]^T = [7, 5, 3, 6]^T$, $V(X_2) = [V_{12}, V_{22}, V_{32}, V_{42}]^T = [5, 2, 9, 4]^T$ and $V(X_3). = [V_{13}, V_{23}, V_{33}, V_{43}]^T = [5, 2, 3, 4]^T$. Applying formula (6) for $i = 1, 2, 3, 4$ one gets finally the following generalized expected values using vectorial notation:

$$[C_1, C_2, C_3, C_4]^T = \sum_{k=1}^{r=3} m(X_k) \cdot V(X_k) = [6.2, 3.8, 4.8, 5.2]^T$$

According to these values, the best alternative to take is A_1 since it has the highest generalized expected payoff.

• Using optimistic attitude, one takes the max value of each row, and applying OWA on each row of $M(X_k)$ for $k = 1$ to r, one gets: $V(X_1) = [V_{11}, V_{21}, V_{31}, V_{41}]^T = [13, 12, 10, 15]^T$, $V(X_2) = [V_{12}, V_{22}, V_{32}, V_{42}]^T = [6, 10, 13, 9]^T$, and $V(X_3) = [V_{13}, V_{23}, V_{33}, V_{43}]^T = [13, 12, 13, 15]^T$. One finally gets $[C_1, C_2, C_3, C_4]^T = [10.9, 11.4, 11.2, 13.2]^T$ and the best alternative to take with optimistic attitude is A_4 since it has the highest generalized expected payoff. Then we have expected payoff as

$$E[C] = \begin{bmatrix} E[C_1] \\ E[C_2] \\ E[C_3] \\ E[C_4] \end{bmatrix} \subset \begin{bmatrix} [6.2; 10.9] \\ [3.8; 11.4] \\ [4.8; 11.2] \\ [5.2; 13.2] \end{bmatrix}$$

2) Implementation of COWA-ER: Let's describe in details each step of COWA-ER. In step 1, we divide each bound of intervals by the max of the bounds to get a new normalized imprecise expected payoff vector $E^{Imp}[C]$. In our example, one gets:

$$E^{Imp}[C] = \begin{bmatrix} [6.2/13.2; 10.9/13.2] \\ [3.8/13.2; 11.4/13.2] \\ [4.8/13.2; 11.2/13.2] \\ [5.2/13.2; 13.2/13.2] \end{bmatrix} \approx \begin{bmatrix} [0.47; 0.82] \\ [0.29; 0.86] \\ [0.36; 0.85] \\ [0.39; 1.00] \end{bmatrix}$$

In step 2, we convert each imprecise value into its bba according to a very natural and simple transformation [17]. Here, we need to consider the finite set of alternatives $\Theta =$

$\{A_1, A_2, A_3, A_4\}$ as FOD. The sources of belief associated with them are obtained from the normalized imprecise expected payoff vector $E^{Imp}[C]$. The modeling for computing a bba associated to the hypothesis A_i from any imprecise value $[a; b] \subseteq [0; 1]$ is very simple and is done as in (7). where $\bar{A}_i$ is the complement of A_i in Θ. With such a simple conversion, one sees that $Bel(A_i) = a$, $Pl(A_i) = b$. The uncertainty is represented by the length of the interval $[a; b]$ and it corresponds to the imprecision of the variable (here the expected payoff) on which the belief function for A_i is defined. In the example, one gets:

TABLE I
BBA'S OF THE ALTERNATIVES USED IN COWA-ER.

Alternatives A_i	$m_i(A_i)$	$m_i(\bar{A}_i)$	$m_i(A_i \cup \bar{A}_i)$
A_1	0.47	0.18	0.35
A_2	0.29	0.14	0.57
A_3	0.36	0.15	0.49
A_4	0.39	0	0.61

In step 3, we use Dempster's rule of combination to obtain[3] the combined bba, which is listed in Table II.

TABLE II
FUSION OF 4 BBA'S WITH DEMPSTER'S RULE FOR COWA-ER.

Focal Element	$m_{Dempster}(.)$
A_1	0.2522
A_2	0.1151
A_3	0.1627
A_4	0.1894
$A_1 \cup A_4$	0.0087
$A_2 \cup A_4$	0.0180
$A_3 \cup A_4$	0.0137
$A_1 \cup A_2 \cup A_4$	0.0368
$A_1 \cup A_3 \cup A_4$	0.0279
$A_2 \cup A_3 \cup A_4$	0.0576
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.1179

In step 4, we use Pignistic Transformation to obtain the bba's corresponding pignistic probability listed in Table III. More efficient (but complex) transformations, like DSmP, could be used instead [19]. Based on the pignistic probability obtained, the decision result is A_1.

TABLE III
PIGNISTIC PROBABILITY BASED ON COWA-ER.

Focal Element	$BetP(.)$
A_1	0.3076
A_2	0.1851
A_3	0.2275
A_4	0.2798

3) Implementation of FCOWA-ER: In step 1 of FCOWA-ER, we normalize each column in $E[C]$, respectively. In our example, one gets:

$$E^{Fuzzy}[C] = \begin{bmatrix} [6.2/6.2; 10.9/13.2] \\ [3.8/6.2; 11.4/13.2] \\ [4.8/6.2; 11.2/13.2] \\ [5.2/6.2; 13.2/13.2] \end{bmatrix} \approx \begin{bmatrix} [1.0000; 0.8258] \\ [0.6129; 0.8636] \\ [0.7742; 0.8485] \\ [0.8387; 1.0000] \end{bmatrix}$$

[3]Other combination rules can be used instead to circumvent the limitations of Dempsters rule discuseed in [19], [20].

Then we obtain two FMFs, which are

$\mu_1 = [1, 0.6129, 0.7742, 0.8387]$;

$\mu_2 = [0.8258, 0.8636, 0.8485, 1.0000]$.

In step 2, by using α-cut approach, μ_1 and μ_2 are converted into two bba's $m_{Pess}(.)$ and $m_{Opti}(.)$ as listed in Table IV. In Step 3, we use Dempster's rule[4] to combine $m_{Pess}(.)$ and

TABLE IV
THE TWO BBA'S TO COMBINE OBTAINED FROM FMFS.

Focal Element	$m_{Pess}(.)$	Focal Element	$m_{Opti}(.)$
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.6129	$A_1 \cup A_2 \cup A_3 \cup A_4$	0.8257
$A_1 \cup A_3 \cup A_4$	0.1613	$A_2 \cup A_3 \cup A_4$	0.0227
$A_1 \cup A_4$	0.0645	$A_2 \cup A_4$	0.0152
A_1	0.1613	A_4	0.1364

$m_{Opti}(.)$ to get $m_{Dempster}(.)$ as listed in Table V.

TABLE V
FUSION OF TWO BBA'S WITH DEMPSTER'S RULE FOR FCOWA-ER.

Focal Element	$m_{Dempster}(.)$
A_1	0.1370
A_4	0.1227
$A_1 \cup A_4$	0.0549
$A_2 \cup A_4$	0.0096
$A_3 \cup A_4$	0.0038
$A_1 \cup A_3 \cup A_4$	0.1370
$A_2 \cup A_3 \cup A_4$	0.0143
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.5207

In step 4, we use again the Pignistic Transformation to get the pignistic probabilities listed in Table VI. Based on

TABLE VI
PIGNISTIC PROBABILITY BASED ON FCOWA-ER.

Focal Element	$BetP(.)$
A_1	0.3403
A_2	0.1397
A_3	0.1826
A_4	0.3374

these probabilities, the decision result is also A_1. The decision results of COWA-ER and FCOWA-ER are the same.

IV. ANALYSES ON FCOWA-ER

A. On computational complexity

In FCOWA-ER, only two bba's are involved in the combination. That is to say only one combination step is needed. Whereas in the original COWA-ER, if there exists q alternatives, there should be $q - 1$ evidence combination operations to do. Furthermore, the bba's obtained in the FCOWA-ER by using α-cut are consonant support (nested in order). This will bring less computational complexity when compared with the bba's generated in the original COWA-ER. In summary, it is clear that the new proposed FCOWA-ER has lower computational complexity.

[4]In fact, and more generally the choice of a rule of combination is entirely left to the preference of the user in our FCOWA-ER methodology.

B. On physical meaning

In this new FCOWA-ER approach, the two different information sources are pessimistic OWA and optimistic OWA. The combination result can be regarded as a tradeoff between these two attitudes. the physical or practical meaning is relatively clear. Furthermore, if the decision-maker has preference on pessimistic or optimistic attitude, one can use discounting in evidence combination to satisfy one's preference. We can set the preference of pessimistic attitude as λ_p and set the the preference of optimistic attitude as λ_o. Then the discounting factor can be obtained as

$$\begin{cases} \beta = \lambda_o/\lambda_p, & \lambda_o \leq \lambda_p \\ \beta = \lambda_p/\lambda_o, & \lambda_p < \lambda_o \end{cases} \tag{10}$$

Then according to the discounting method [16], one will take:

$$\begin{cases} m_\beta(X) = \beta \cdot m(X), \quad \text{for } X \neq \Theta \\ m_\beta(\Theta) = \beta \cdot m(\Theta) + (1-\beta) \end{cases} \tag{11}$$

If $\lambda_o \leq \lambda_p$, then $m(.)$ in (11) should be $m_{Opti}(.)$; If $\lambda_p \leq \lambda_o$, then $m(.)$ in (11) should be $m_{Pess}(.)$; By using the discounting and choosing a combination rule, the decision maker's has a flexibility in his decision-making process.

C. On management of uncertainty

In the FCOWA-ER, we first define the bba vertically taking into account the uncertainty between alternatives for the pessimist attitude and for the optimistic attitude. Then we combine two columns. The uncertainty incorporated in the FMF obtained, which represents the possibility of each alternative to be chosen as the final decision result. Based on α-cut approach, the FMF is transformed into bba. The uncertainty is thus transformed to the bba. Although based on each column, only the information of pessimistic or optimistic is used, the combination operation followed can use both the two information sources (pessimistic and optimistic attitudes). Thus the available information can be fully used in FCOWA-ER. In COWA-ER, the modeling for each row (interval) takes into account the true uncertainty one has on the bounds of payoff for each alternative, then after modeling each bba $m_i(.)$, one combines them "vertically" to take into account the uncertainty between alternatives.

Although the ways to manage the uncertainty incorporated in are different for COWA-ER and FCOWA-ER, they both utilize (differently) the whole available information.

D. On robustness to error scoring

Based on many experiments, we have observed that almost all the decision results given by FCOWA-ER agree[5] with COWA-ER results and are rational. However when the difference among the values in payoff matrix is significant, the COWA-ER can yield to wrong decisions whereas FCOWA-ER yields to rational decisions as illustrated in Example 2 below.

[5]when using the same rule of combination in steps 3, and the same probabilistic transformation in steps 4.

Example 2 Let's take states $S = \{S_1, S_2, S_3, S_4, S_5\}$ with associated bba $m(S_1 \cup S_2 \cup S_3 \cup S_4 \cup S_5) = 1$, and consider alternatives $A = \{A_1, A_2, A_3, A_4\}$ and the payoffs matrix:

$$C = \begin{bmatrix} 12 & 11 & 10 & 120 & 7 \\ 9 & 10 & 6 & 110 & 3 \\ 7 & 13 & 5 & 100 & 6 \\ 6 & 2 & 3 & 150 & 4 \end{bmatrix} \tag{12}$$

We see that the difference between max value and min value of each line is significant. For example, in the fourth row of C, only S_4 brings extremely high score for A_4 whereas other states bring homogeneous low score values for A_4. Whatever state of nature we consider S_1, S_2,..., or S_5, A_1 is either the top 1 or top 2 choice according to the ranks of the alternatives for states S_i, $i = 1, \ldots, 5$ as shown below:

	A_1	A_2	A_3	A_4
S_1	1	2	3	4
S_2	2	3	1	4
S_3	1	2	3	4
S_4	2	3	4	1
S_5	1	4	2	3

So, intuitively, according to the principle of majority voting, the decision result should be A_1 but not A_4. According to rank-level fusion, the decision result should also be A_1.

The expected payoffs are:

$$E[C] = \begin{bmatrix} [7, & 120] \\ [3, & 110] \\ [5, & 100] \\ [2, & 150] \end{bmatrix}$$

- Using COWA-ER, one has

$$E^{Imp}[C] = \begin{bmatrix} [0.0467, & 0.8000] \\ [0.0200, & 0.7333] \\ [0.0333, & 0.6667] \\ [0.0133, & 1.0000] \end{bmatrix}$$

The bba's to combine are listed in Table VII and the combination results by using Dempster's rule are in Table VIII.

TABLE VII
EXAMPLE 2: BBA'S OF THE ALTERNATIVES USED IN COWA-ER.

Alternatives A_i	$m_i(A_i)$	$m_i(\bar{A}_i)$	$m_i(A_i \cup \bar{A}_i)$
A_1	0.0467	0.2000	0.7533
A_2	0.0200	0.2667	0.7133
A_3	0.0333	0.3333	0.6334
A_4	0.0133	0	0.9867

The pignistic probabilities listed in IX indicate that the decision result[6] of COWA-ER is A_4.

- Using FCOWA-ER, one has

$$E^{Fuzzy}[C] = \begin{bmatrix} [1.0000, & 0.8000] \\ [0.4286, & 0.7333] \\ [0.7143, & 0.6667] \\ [0.2857, & 1.0000] \end{bmatrix}$$

In FCOWA-ER, the bba's to combine are listed in Table X and their Dempster's combination is listed in Table XI.

TABLE VIII
EXAMPLE 2: DEMPSTER'S FUSION OF 4 BBA'S FOR COWA-ER.

Focal Element	$m_{Dempster}(.)$
A_1	0.0438
A_2	0.0182
A_3	0.0309
A_4	0.0297
$A_1 \cup A_4$	0.0664
$A_2 \cup A_4$	0.0471
$A_3 \cup A_4$	0.0335
$A_1 \cup A_2 \cup A_4$	0.1775
$A_1 \cup A_3 \cup A_4$	0.1261
$A_2 \cup A_3 \cup A_4$	0.0895
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.3373

TABLE IX
EXAMPLE 2: PIGNISTIC PROBABILITY BASED ON COWA-ER.

Focal Element	$BetP(.)$
A_1	0.2625
A_2	0.2152
A_3	0.2038
A_4	0.3185

TABLE X
EXAMPLE 2: BBA'S OF THE ALTERNATIVES USED IN FCOWA-ER.

Focal Element	$m_{Pess}(.)$	Focal Element	$m_{Opti}(.)$
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.2857	$A_1 \cup A_2 \cup A_3 \cup A_4$	0.6667
$A_1 \cup A_2 \cup A_3$	0.1429	$A_1 \cup A_3 \cup A_4$	0.0667
$A_1 \cup A_3$	0.2857	$A_3 \cup A_4$	0.0667
A_1	0.2857	A_1	0.1999

TABLE XI
EXAMPLE 2: DEMPSTER'S FUSION OF THE TWO BBA'S FOR FCOWA-ER.

Focal Element	$m_{Dempster}(.)$
A_1	0.3223
A_4	0.0667
$A_1 \cup A_2$	0.0111
$A_1 \cup A_3$	0.2222
$A_1 \cup A_4$	0.0222
$A_1 \cup A_2 \cup A_3$	0.1111
$A_1 \cup A_2 \cup A_4$	0.0222
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.2222

The pignistic transformation of $m_{Dempster}(.)$ yields to the pignistic probabilities listed in Table XII.

TABLE XII
EXAMPLE 2: PIGNISTIC PROBABILITY BASED ON FCOWA-ER.

Focal Element	$BetP(.)$
A_1	0.5500
A_2	0.1056
A_3	0.2037
A_4	0.1407

Based on the pignistic probabilities, the decision result obtained with FCOWA-ER is now A_1, which is the correct one. In this example, FCOWA-ER shows its robustness when compared with COWA-ER.

E. On the normalization procedures

In fact, there exist at least three normalization procedures that we briefly recall below. Suppose $\mathbf{x}$ is the original vector input, $\mathbf{x_i}$ represents the ith dimension of $\mathbf{x}$. y_i represents the ith dimension of the normalized vector $\mathbf{y}$, then we examine the following three types of normalization:

1) Type I: $y_i = x_i / \max(\mathbf{x})$
2) Type II: $y_i = (x_i - \min(\mathbf{x}))/(\max(\mathbf{x}) - \min(\mathbf{x}))$
3) Type III: $y_i = x_i / \sum_j (x_j)$

To verify wether the decision results of COWA-ER and the new FCOWA-ER can be affected by the normalization procedure, we did some tests as follows. We randomly generate payoff matrices and use all the three types of normalization approaches in COWA-ER and FCOWA-ER respectively. Then we make comparisons among the results obtained. We repeat the experiment 50 times (50 Monte-Carlo runs). Based on our simulation results, we find that normalization approaches can affect the decision results of COWA-ER and FCOWA-ER, although the ratio of disagreement among different normalization approach is small (about 1 to 2 times of disagreement out of 50 experiments in average). Example 3 is a case where the disagreement occurs due to the different types of normalization.

Example 3: We consider the following payoff matrix

$$C = \begin{bmatrix} 15 & 5 & 30 \\ 5 & 40 & 40 \\ 40 & 30 & 30 \\ 15 & 10 & 40 \end{bmatrix}$$

The bba is $m(X_1) = 0.5439$, $m(X_2) = 0.3711$, $m(X_3) = 0.0849$, where $X_1 = S_2 \cup S_3$, $X_2 = S_1 \cup S_2$ and $X_3 = S_1 \cup S_3$.

• Using COWA-ER, based on normalization Type I, Type II and Type III, we can obtain the corresponding expected payoffs

$$E_I[C] = \begin{bmatrix} [0.1462, & 0.6108] \\ [0.6009, & 1.0000] \\ [0.7500, & 0.8640] \\ [0.2606, & 0.7680] \end{bmatrix}$$

$$E_{II}[C] = \begin{bmatrix} [0.0000, & 0.5442] \\ [0.5326, & 1.0000] \\ [0.7072, & 0.8407] \\ [0.1340, & 0.7283] \end{bmatrix}$$

$$E_{III}[C] = \begin{bmatrix} [0.0292, & 0.1221] \\ [0.1202, & 0.2000] \\ [0.1500, & 0.1728] \\ [0.0521, & 0.1536] \end{bmatrix}$$

Then we obtain the pignistic probabilities listed in Table XIII. From Table XIII, one sees that the decision result with Type III normalization is A_2 while those of Type I and Type II yields A_3.

TABLE XIII
EXAMPEL 3: PIGNISTIC PROB. FOR TYPES I, II & III AND COWA-ER.

Focal Element	$BetP_I(.)$	$BetP_{II}(.)$	$BetP_{III}(.)$
A_1	0.0587	0.0388	0.1690
A_2	0.3203	0.3324	0.3180
A_3	0.5223	0.5444	0.2920
A_4	0.0987	0.0844	0.2210

[6]based on max of $BetP(.)$.

• Using FCOWA-ER, based on normalization of Type I, Type II and Type III, we get the corresponding expected payoffs

$$E_I[C] = \begin{bmatrix} [0.1950, & 0.6108] \\ [0.8013, & 1.0000] \\ [1.0000, & 0.8640] \\ [0.3475, & 0.7680] \end{bmatrix}$$

$$E_{II}[C] = \begin{bmatrix} [0.0000, & 0.0000] \\ [0.7531, & 1.0000] \\ [1.0000, & 0.6506] \\ [0.1894, & 1.4040] \end{bmatrix}$$

$$E_{III}[C] = \begin{bmatrix} [0.0832, & 0.1884] \\ [0.3419, & 0.3084] \\ [0.4267, & 0.2664] \\ [0.1483, & 1.2368] \end{bmatrix}$$

Then we can get the pignistic probabilities listed in Table XIV. From Table XIV, one sees that the decision with Type II normalization is A_2 while those of Type I and Type III yields A_3.

TABLE XIV
EXAMPLE 3: PIGNISTIC PROBABILITY BASED ON FCOWA-ER

Focal Element	$BetP_I(.)$	$BetP_{II}(.)$	$BetP_{III}(.)$
A_1	0.0306	0.0000	0.0306
A_2	0.4118	0.5421	0.4118
A_3	0.4763	0.4300	0.4763
A_4	0.0813	0.0279	0.0813

So in a little percentage of cases, we must be cautious when choosing a normalization procedure and so far there is no clear theoretical answer for the choice of the most adapted normalization procedure. We prefer the Type I normalization procedure since it is very simple and intuitively appealing.

V. CONCLUSION

In this paper, we have proposed a fuzzy cautious OWA method using evidential reasoning (FCOWA-ER) to implement the multi-criteria decision making, where evidence theory, fuzzy membership functions and OWA are used jointly. This method has less computational complexity and has clearer physical meaning. Furthermore, it is more robust to the error scoring in MCDM. Experimental results and related analyses show that our FCOWA-ER is interesting and flexible because its three main specifications can be adapted easily for working: 1) with other rules of combination than Dempster's rule, 2) with other probabilistic transformations than BetP, and 3) with different normalization procedures. Of course the performances of FCOWA-ER depend on the choice of these three main specifications taken by the MCDM system designer. The method to generate the bba from the FMF based on α-cut depends on the selection of the parameter vector of α. The impact of the choice of the specifications as well as α to evaluate the performance of FCOWA-ER will be further analyzed in our future works. This paper was devoted to the theoretical developemnt of FCOWA-ER and its evaluation for applications to real MCDM problems is part of our future research works.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China (No.61104214, No.67114022, No. 61074176, No. 60904099), Fundamental Research Funds for the Central Universities, China Postdoctoral Science Foundation (No.20100481337, No.201104670), Research Fund of Shaanxi Key Laboratory of Electronic Information System Integration (No.201101Y17), Chongqing Natural Science Foundation (No. CSCT, 2010BA2003) and Aviation Science Foundation (No. 2011ZC53041).

REFERENCES

[1] B. Roy, "Paradigms and challenges", in Multiple Criteria Decision Analysis : State of the art surveys, Vol.78 of Int. Series in Oper. Research and& Management Sci. (Chap. 1), pp. 1–24, Springer, 2005.
[2] R. Yager, "On ordered weighted averaging operators in multi-criteria decision making", *IEEE Transactions on Systems, Man and Cybernetics*, 18:183–190, 1988.
[3] R. Yager, J. Kacprzyk, *The ordered weighted averaging operator: theory and applications*, Kluwer Academic Publishers, MA: Boston, 1997.
[4] M. O'Hagan, "Aggregating template rule antecedents in real-time expert systems with fuzzy set logic", Proc. of 22nd Ann. IEEE Asilomar Conf. on Signals, Systems and Computers, Pacific Grove, CA, 1988, 681–689.
[5] D. Filev, R. Yager, "On the issue of obtaining OWA operator weights", *Fuzzy Sets and Systems*, 94: 157–169, 1998.
[6] B.S. Ahn, "Parameterized OWA operator weights: an extreme point approach", *International Journal of Approximate Reasoning*, 51: 820–831, 2010.
[7] R. Yager, "Using stress functions to obtain OWA operators", *IEEE Transactions on Fuzzy Systems*, 15: 1122C-1129, 2007.
[8] Y.M. Wang, C. Parkan, "A minimax disparity approach obtaining OWA operator weights", *Information Sciences*, 175: 20–29, 2005.
[9] A. Emrouznejad, G.R. Amin, "Improving minimax disparity model to determine the OWA operator weights", *Information Sciences*, 180: 1477–1485, 2010.
[10] Y.M. Wang, Y. Luo, Z. Hua, "Aggregating preference rankings using OWA operator weights", *Information Sciences* 177: 3356–3363, 2007.
[11] M. Beynon, "A method of aggregation in DS/AHP for group decision-making with non-equivalent importance of individuals in the group", *Computers and Operations Research*, 32: 1881–1896, 2005.
[12] J. Dezert, J.-M. Tacnet, M. Batton-Hubert, F. Smarandache,"Multi-criteria decision making based on DSmT-AHP", Proc. of Belief 2010 Int. Workshop, Brest, France, 1-2 April, 2010.
[13] J.-M. Tacnet, M. Batton-Hubert, J. Dezert, "A two-step fusion process for multi-criteria decision applied to natural hazards in mountains", Proc. of Belief 2010 Int. Workshop, Brest, France, 2010.
[14] R. Yager, "Decision making under Dempster-Shafer uncertainties", Studies in Fuzziness and Soft Computing, 219:619–632, 2008.
[15] J.-M. Tacnet, J. Dezert, " Cautious OWA and evidential reasoning for decision making under uncertainty", Proc. of 14th International Conference on Information Fusion Chicago, Illinois, USA, July 5-8, 2011, 2074–2081.
[16] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University, Princeton, 1976.
[17] J. Dezert,"Autonomous navigation with uncertain reference points using the PDAF", *Multitarget-Multisensor Tracking*, Vol 2, pp 271–324, Y. Bar-Shalom Editor, Artech House, 1991.
[18] M. C. Florea, A.-L. Jousselme, D. Grenier, É. Bossé, "Approximation techniques for the transformation of fuzzy sets into random sets", *Fuzzy Sets and Systems*, 159: 270–288, 2008.
[19] F. Smarandache, J. Dezert, "Advances and applications of DSmT for information fusion (Collected works)", Vol. 1-3, American Research Press, 2004–2009. http://www.gallup.unm.edu/~smarandache/DSmT.htm
[20] J. Dezert, P. Wang, A. Tchamova, "On the validity of Dempster-Shafer Theory", in Proceedings of Fusion 2012 Conf., Singapore, July 2012.

Soft ELECTRE TRI Outranking Method Based on Belief Functions

Jean Dezert
Jean-Marc Tacnet

Originally published as Dezert J., Tacnet J.-M., *Soft ELECTRE TRI outranking method based on belief functions*, Proc. Of Fusion 2012, Singapore, July 2012, and reprinted with permission.

***Abstract*—Main decisions problems can be described into choice, ranking or sorting of a set of alternatives. The classical ELECTRE TRI (ET) method is a multicriteria-based outranking sorting method which allows to assign alternatives into a set of predetermined categories. ET method deals with situations where indifference is not transitive and solutions can sometimes appear uncomparable. ET suffers from two main drawbacks: 1) it requires an arbitrary choice of λ-cut step to perform the outranking of alternatives versus profiles of categories, and 2) an arbitrary choice of attitude for final assignment of alternatives into the categories. ET finally gives a final binary (hard) assignment of alternatives into categories. In this paper we develop a soft version of ET method based on belief functions which circumvents the aforementioned drawbacks of ET and allows to obtain both a soft (probabilistic) assignment of alternatives into categories and an indicator of the consistency of the soft solution. This Soft-ET approach is applied on a concrete example to show how it works and to compare it with the classical ET method.**

Keywords: ELECTRE TRI, information fusion, belief functions, outranking methods, multicriteria analysis.

I. Introduction

Multi-criteria decision analysis aims to choose, sort or rank alternatives or solutions according to criteria involved in the decision-making process. The main steps of a multi-criteria analysis consist in identifying decision purposes, defining criteria, eliciting preferences between criteria, evaluating alternatives or solutions and analyzing sensitivity with regard to weights, thresholds, etc. A difference has to be done between total and partial aggregation methods:

- Total aggregation methods such as the Multi-Attribute Utility Theory (M.A.U.T.) [1], [2] synthesizes in a unique value the partial utility related to each criterion and chosen by the decision-maker (DM). Each partial utility function transforms any quantitative evaluation of criterion into an utility value. The additive method is the simplest method to aggregate those utilities.
- Partial aggregation methods which are not based on the principle of preference transitivity. The ELECTRE TRI (ET) outranking method inspired by Roy [3] and finalized by Yu in [4] belongs to this family and it is the support of the research work presented in this paper.

ELECTRE TRI (electre tree) is an evolution of the ELECTRE methods introduced in 1960's by Roy [5] which remain widespread methods used in operational research. The acronym ELECTRE stands for "ELimination Et Choix Traduisant la REalité (Elimination and Choice Expressing the Reality). ET is simpler and more general than the previous ELECTRE methods which have specificities given their context of applications. A good introduction to ET methods with substantial references and detailed historical survey can be found in [6] and additional references in [7]. This paper proposes a methodology inspired by the ET method able to help decision based on imperfect information for soft assignment of alternatives into a given set of categories defined by predeterminate profiles. Our method, called "Soft ELECTRE TRI" (or just SET for short), is based on belief functions. It allows to circumvent the problem of arbitrary choice of λ-cut of the outranking step of ET, and the ad-choice of attitude in the final assignment step of ET as well. Contrariwise to ET which solves the hard assignment problem, SET proposes a new solution for solving the assignment problem in a soft manner. This paper is organized as follows. In Section II, we recall the principles of ET method with its main steps. In Section III, we present in details our new SET method with emphasize on its differences with classical ET. In Section IV, we apply ET and SET on a concrete example proposed by Maystre [8] to show how they work and to make a comparison between the two approaches. Section V concludes this paper and proposes some perspectives of this work.

II. The ELECTRE TRI (ET) method

Outranking methods like the ET method presented in this section are relevant for Multi-Criteria Decision Analysis (MCDA) [6] when:

- alternatives are evaluated on an ordinal scale;
- criteria are strongly heterogeneous by nature (e.g. comfort, price, pollution);
- compensation of the loss on one criterion by a gain on an another is unacceptable;
- small differences of evaluations are not individually significant while the accumulation of several of these differences may become significant.

We are concerned with an assignment problem in complex situations where several given alternatives have to be assigned to known categories based on multiple criteria. The categories

are defined by profiles values (bounds) for each criteria involved in the problem under consideration as depicted in Fig. 1 below.

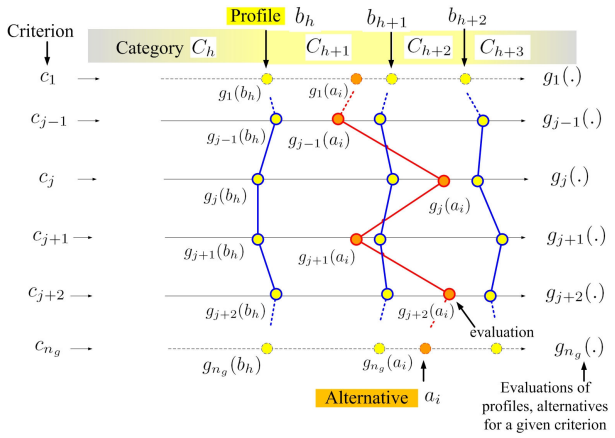


Figure 1: ET aims to assign a category to alternatives.

The ET method is a multicriteria-based outranking sorting method proposing a hard assignment of alternatives a_i in categories C_h. More precisely, the alternatives $a_i \in \mathbf{A}$, $i = 1, \ldots, n_a$ are committed to ordered categories $C_h \in \mathbf{C}$, $h = 1, \ldots, n_h$ according to criteria c_j, $j \in \mathbf{J} = \{1, \ldots, n_g\}$. Each category C_h is delimited by the set of its lower and upper limits b_{h-1} and b_h with respect to their evaluations $g_j(b_{h-1})$ and $g_j(b_h)$ for each criterion c_j ($g_j(.)$ represents the evaluations of alternatives, profiles for a given criterion c_j). By convention, $b_0 \leq b_1 \ldots \leq b_{n_h}$. b_0 is the lower (minimal) profile bound and b_{n_h} is the upper (maximal) profile bound. The overall profile b_h is defined through the set of values $\{g_1(b_h), g_2(b_h), \ldots, g_{n_g}(b_h)\}$ represented by the vertical lines joining the yellow dots in Fig. 1. The outranking relations are based on the calculation of partial concordance and discordance indices from which global concordance and credibility indices [4], [9] are derived based on an arbitrary λ-cut strategy. The final assignment (sorting procedure) of alternatives to categories operated by ET is a hard (binary) assignment based on an arbitrary selected attitude choice (optimistic or pessimistic). ET method can be summarized by the following steps:

- ET-Step 1: Computation of partial concordance indices $c_j(a_i, b_h)$ and $c_j(b_h, a_i)$), and partial discordances indices $d_j(a_i, b_h)$ and $d_j(b_h, a_i)$);
- ET-Step 2: Computation of the global (overall) concordance indices $c(a_i, b_h)$ and $c(b_h, a_i)$ to obtain credibility indices $\rho(a_i, b_h)$ and $\rho(b_h, a_i)$;
- ET-Step 3: Computation of the fuzzy outranking relation grounded on the credibility indices $\rho(a_i, b_h)$ and $\rho(b_h, a_i)$; and apply a λ-cut to get the crisp outranking relation;
- ET-Step 4: Final hard (binary) assignment of a_i into C_h is based on the crisp outranking relation and in adopting either a pessimistic (conjunctive), or an optimistic (disjunctive) attitude.

Let's explain a bit more in details the steps of ET and the computation of the indices necessary for the implementation of the ET method.

A. ET-Step 1: Partial indices

In ET method, the partial concordance index $c_j(a_i, b_h)$ (resp. $c_j(b_h, a_i)$) expresses to which extent the evaluations of a_i and b_h (respectively b_h and a_i)) are concordant with the assertion *"a_i is at least as good as b_h"* (respectively *"b_h is at least as good as a_i"*). $c_j(a_i, b_h) \in [0, 1]$, based on a given criterion $g_j(.)$, is computed from the difference of the criterion evaluated for the profile b_h, and the same criterion evaluated for the alternative a_i. If the difference[1] $g_j(b_h) - g_j(a_i)$ is less (or equal) to a given indifference threshold $q_j(b_h)$ then a_i and b_h are considered indifferent based on the criterion $g_j(.)$. If the difference $g_j(b_h) - g_j(a_i)$ is strictly greater to given preference threshold $p_j(b_h)$ then a_i and b_h are considered different based on $g_j(.)$. When $g_j(b_h) - g_j(a_i) \in [q_j(b_h), p_j(b_h)]$, the partial concordance index $c_j(a_i, b_h)$ is computed from a linear interpolation corresponding to a weak difference. Mathematically, the partial concordance indices $c_j(a_i, b_h)$ and $c_j(b_h, a_i)$ are obtained by:

$$c_j(a_i, b_h) \triangleq \begin{cases} 0 & \text{if} \quad g_j(b_h) - g_j(a_i) \geq p_j(b_h) \\ 1 & \text{if} \quad g_j(b_h) - g_j(a_i) < q_j(b_h) \\ \frac{g_j(a_i) + p_j(b_h) - g_j(b_h)}{p_j(b_h) - q_j(b_h)} & \text{otherwise} \end{cases} \tag{1}$$

and

$$c_j(b_h, a_i) \triangleq \begin{cases} 0 & \text{if} \quad g_j(a_i) - g_j(b_h) \geq p_j(b_h) \\ 1 & \text{if} \quad g_j(a_i) - g_j(b_h) < q_j(b_h) \\ \frac{g_j(b_h) + p_j(b_h) - g_j(a_i)}{p_j(b_h) - q_j(b_h)} & \text{otherwise.} \end{cases} \tag{2}$$

The partial discordance index $d_j(a_i, b_h)$ (resp. $d_j(b_h, a_i)$) expresses to which extent the evaluations of a_i and b_h (resp. b_h and a_i) is opposed to the assertion *"a_i is at least as good as b_h"* (resp.*"b_h is at least as good as a_i"*). These indices depend on a possible veto condition expressed by the choice of a veto threshold $v_j(b_h)$ (such as $v_j(b_h) \geq p_j(b_h) \geq q_j(b_h) \geq 0$) imposed on some criterion $g_j(.)$. They are defined by [4], [9]:

$$d_j(a_i, b_h) \triangleq \begin{cases} 0 & \text{if} \quad g_j(b_h) - g_j(a_i) < p_j(b_h) \\ 1 & \text{if} \quad g_j(b_h) - g_j(a_i) \geq v_j(b_h) \\ \frac{g_j(b_h) - g_j(a_i) - p_j(b_h)}{v_j(b_h) - p_j(b_h)} & \text{otherwise} \end{cases} \tag{3}$$

and

$$d_j(b_h, a_i) \triangleq \begin{cases} 0 & \text{if} \quad g_j(a_i) - g_j(b_h) \leq p_j(b_h) \\ 1 & \text{if} \quad g_j(a_i) - g_j(b_h) > v_j(b_h) \\ \frac{g_j(a_i) - g_j(b_h) - p_j(b_h)}{v_j(b_h) - p_j(b_h)} & \text{otherwise.} \end{cases} \tag{4}$$

[1]For convenience, we assume here an increasing preference order. A decreasing preference order [9] can be managed similarly by multiplying criterion values by -1.

B. ET-Step 2: Global concordance and credibility indices

- **The global concordance indices**: The global concordance index $c(a_i, b_h)$ (respectively $c(b_h, a_i)$) expresses to which extent the evaluations of a_i and b_h on all criteria (respectively b_h and a_i) are concordant with the assertions *"a_i outranks b_h"* (respectively *"b_h outranks a_i"*). In ET method, $c(a_i, b_h)$ (resp. $c(b_h, a_i)$) is computed by the weighted average of partial concordance indices $c_j(a_i, b_h)$ (resp. $c_j(b_h, a_i)$). That is

$$c(a_i, b_h) = \sum_{j=1}^{n_g} w_j c_j(a_i, b_h) \tag{5}$$

and

$$c(b_h, a_i) = \sum_{j=1}^{n_g} w_j c_j(b_h, a_i) \tag{6}$$

where the weights $w_j \in [0, 1]$ represent the relative importance of each criterion $g_j(.)$ in the evaluation of the global concordance indices. The weights add to one. Since all $c_j(a_i, b_h)$ and $c_j(a_i, b_h)$ belong to $[0; 1]$, $c(a_i, b_h)$ and $c(b_h, a_i)$ given by (5) and (6) also belong to $[0; 1]$.

- **The global credibility indices**: The degree of credibility of the outranking relation denoted as $\rho(a_i, b_h)$ (respectively $\rho(b_h, a_i)$) expresses to which extent *"a_i outranks b_h"* (respectively *"b_h outranks a_i"*) according to the global concordance index $c(a_i, b_h)$ and the discordance indices $d_j(a_i, b_h)$ for all criteria (respectively $c(b_h, a_i)$ and $d_j(b_h, a_i)$). In ET method, these credibility indices $\rho(a_i, b_h)$ (resp. $\rho(b_h, a_i)$) are computed by discounting (weakening) the global concordance indices $c(a_i, b_h)$ given by (5) (resp. $c(b_h, a_i)$ given by (6)) by a discounting factor $\alpha(a_i, b_h)$ in $[0; 1]$ (resp. $\alpha(b_h, a_i)$) as follows:

$$\begin{cases} \rho(a_i, b_h) = c(a_i, b_h)\alpha(a_i, b_h) \\ \rho(b_h, a_i) = c(b_h, a_i)\alpha(b_h, a_i) \end{cases} \tag{7}$$

The discounting factors $\alpha(a_i, b_h)$ and $\alpha(b_h, a_i)$ are defined by [9], [10]:

$$\alpha(a_i, b_h) \triangleq \begin{cases} 1 & \text{if} \quad \mathbf{V}_1 = \emptyset \\ \prod_{j \in \mathbf{V}_1} \frac{1 - d_j(a_i, b_h)}{1 - c(a_i, b_h)} & \text{if} \quad \mathbf{V}_1 \neq \emptyset \end{cases} \tag{8}$$

$$\alpha(b_h, a_i) \triangleq \begin{cases} 1 & \text{if} \quad \mathbf{V}_2 = \emptyset \\ \prod_{j \in \mathbf{V}_2} \frac{1 - d_j(b_h, a_i)}{1 - c(b_h, a_i)} & \text{if} \quad \mathbf{V}_2 \neq \emptyset \end{cases} \tag{9}$$

where $\mathbf{V}_1$ (resp. $\mathbf{V}_2$) is the set of indexes j where the partial discordance indices $d_j(a_i, b_h)$ (reps. $d_j(b_h, a_i)$) is greater than the global concordance index $c(a_i, b_h)$ (resp. $c(b_h, a_i)$), that is:

$$\mathbf{V}_1 \triangleq \{j \in \mathbf{J} | d_j(a_i, b_h) > c(a_i, b_h)\} \tag{10}$$

$$\mathbf{V}_2 \triangleq \{j \in \mathbf{J} | d_j(b_h, a_i) > c(b_h, a_i)\} \tag{11}$$

C. ET-Step 3: Fuzzy and crisp outranking process

Outranking relations result from the transformation of fuzzy outranking relation (corresponding to credibility indices) into a crisp outranking relation[2] S done by means of a λ-cut [9]. λ is called *cutting level*. λ is the smallest value of the credibility index $\rho(a_i, b_h)$ compatible with the assertion "a_i *outranks* b_h". Similarly λ is the smallest value of the credibility index $\rho(b_h, a_i)$ compatible with the assertion "b_h *outranks* a_i". In practice the choice of λ value is not easy and is done arbitrary or based on a sensitivity analysis. More precisely, the crisp outranking relation S is defined by

$$\begin{cases} \rho(a_i, b_h) \geq \lambda \Longrightarrow a_i \ S \ b_h \\ \rho(b_h, a_i) \geq \lambda \Longrightarrow b_h \ S \ a_i \end{cases} \tag{12}$$

Binary relations of preference ($>$), indifference (I), incomparability (R) are defined according to (13):

$$\begin{cases} a_i I b_h \Longleftrightarrow a_i \ S \ b_h \text{ and } b_h \ S \ a_i \\ a_i > b_h \Longleftrightarrow a_i \ S \ b_h \text{ and not } b_h \ S \ a_i \\ a_i < b_h \Longleftrightarrow \text{not } a_i \ S \ b_h \text{ and } b_h \ S \ a_i \\ a_i R b_h \Longleftrightarrow \text{not } a_i \ S \ b_h \text{ and not } b_h \ S \ a_i \end{cases} \tag{13}$$

D. ET-Step 4: Hard assignment procedure

Based on outranking relations between all pairs of alternatives and profiles of categories, two attitudes can be used in ET to assign each alternative a_i into a category C_h [6]. These attitudes yields to a hard assignment solution where each alternative belongs or doesn't belong to a category (binary assignment) and there is no measure of the confidence of the assignment in this last step of ET method. The pessimistic and optimistic hard assignments are realized as follows:

- Pessimistic hard assignment: a_i is compared with b_k, b_{k-1}, b_{k-2}, ..., until a_i outranks b_h where $h \leq k$. The alternative a_i is then assigned to the highest category C_h, that is $a_i \rightarrow C_h$, if $\rho(a_i, b_h) \geq \lambda$.
- Optimistic hard assignment: a_i is compared successively to b_1, b_2, ... b_h, ... until b_h outranks a_i. The alternative a_i is assigned to the lowest category C_h, $a_i \rightarrow C_h$, for which the upper profile b_h is preferred to a_i.

III. The new Soft ELECTRE TRI (SET) method

The objective and motivation of this paper are to improve the appealing ET method in order to provide a soft assignment procedure of alternatives into categories, and to eliminate the drawback concerning both the choice of λ-cut level in ET-Step 3 and the choice of attitude in ET-Step 4. Soft assignment reflects the confidence one has in the assignment which can be a very useful property in applications requiring multi criteria decision analysis. To achieve such purpose and due to long experience in working with belief functions (BF), it has appeared clearly that BF can be very useful for developing a "soft-assignment" version of the classical ET presented in the previous section. We call this new method the "Soft ELECTRE

[2]It is denoted S because Outranking translates to "Surclassement" in French.

TRI" method (SET for short) and we present it in details in this section.

Before going further, it is necessary to recall briefly the definition of a mass of belief $m(.)$ (also called basic belief assignment, or bba), a credibility function Bel(.) and the plausibility function Pl(.) defined over a finite set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ of mutually exhaustive and exclusive hypotheses. Belief functions have been introduced by Shafer in his development of Dempster-Shafer Theory (DST), see [11] for details. In DST, Θ is called the frame of discernment of the problem under consideration. By convention the power-set (i.e. the set of all subsets of Θ) is denoted 2^Θ since its cardinality is $2^{|\Theta|}$. A basic belief assignment provided by a source of evidence is a mapping $m(.) : 2^\Theta \rightarrow [0,1]$ satisfying

$$m(\emptyset) = 0 \qquad \text{and} \qquad \sum_{X \in 2^\Theta} m(X) = 1 \tag{14}$$

The measures of credibility and plausibility of any proposition $X \in 2^\Theta$ are defined from $m(.)$ by

$$\text{Bel}(X) \triangleq \sum_{\substack{Y \subseteq X \\ Y \in 2^\Theta}} m(Y) \tag{15}$$

$$\text{Pl}(X) \triangleq \sum_{\substack{Y \cap X \neq \emptyset \\ Y \in 2^\Theta}} m(Y) \tag{16}$$

$\text{Bel}(X)$ and $\text{Pl}(X)$ are usually interpreted as lower and upper bounds of the unknown probability of X. $U(X) = \text{Pl}(X) - \text{Bel}(X)$ reflects the uncertainty on X. The belief functions are well adapted to model uncertainty expressed by a given source of evidence. For information fusion purposes, many solutions have been proposed in the literature [12] to combine bba's efficiently for pooling evidences arising from several sources.

As for the classical ET method, there are four main steps in our new SET method. However, the SET steps are different from the ET steps. The four steps of SET, that are actually very specific and improves the ET steps, are:

- SET-Step 1: Computation of partial concordance indices $c_j(a_i, b_h)$ and $c_j(b_h, a_i)$, partial discordances indices $d_j(a_i, b_h)$ and $d_j(b_h, a_i)$, and also partial uncertainty indices $u_j(a_i, b_h)$ and $u_j(b_h, a_i)$ thanks to a smooth sigmoidal model for generating bba's [13].
- SET-Step 2: Computation of the global (overall) concordance indices $c(a_i, b_h)$, $c(b_h, a_i)$, discordance indices $d(a_i, b_h)$, $d(b_h, a_i)$, and uncertainty indices $u(a_i, b_h)$, $u(b_h, a_i)$;
- SET-Step 3: Computation of the probabilized outranking relations grounded on the global indices of SET-Step 2. The probabilization is directly obtained and thus eliminates the arbitrary λ-cut strategy necessary in ET.
- SET-Step 4: Final soft assignment of a_i into C_h based on combinatorics of probabilized outranking relations.

Let's explain in details the four steps of SET and the computation of the indices necessary for the implementation of the SET method.

A. SET-Step 1: Partial indices

In SET, a sigmoid model is proposed to replace the original truncated trapezoidal model for computing concordance and discordance indices of the ET method. The sigmoidal model has been presented in details in [13] and is only briefly recalled here. We consider a binary frame of discernment[3] $\Theta \triangleq \{c, \bar{c}\}$ where c means that the alternative a_i is concordant with the assertion "a_i is at least as good as profile b_h", and $\bar{c}$ means that the alternative a_i is opposed (discordant) to this assertion. We can compute a basic belief assignment (bba) $m_{ih}(.)$ defined on 2^Θ for each pair (a_i, b_h). $m_{ih}(.)$ is defined from the combination (fusion) of the local bba's $m^j_{ih}(.)$ evaluated from each possible criteria $g_j(.)$ as follows: $m^j_{ih}(.) = [m_1 \oplus m_2](.)$ is obtained by the fusion[4] (denoted symbolically by $\oplus$) of the two following simple bba's defined by:

focal element	$m_1(.)$	$m_2(.)$
c	$f_{s_c,t_c}(g)$	0
$\bar{c}$	0	$f_{-s_{\bar{c}},t_{\bar{c}}}(g)$
$c \cup \bar{c}$	$1 - f_{s_c,t_c}(g)$	$1 - f_{-s_{\bar{c}},t_{\bar{c}}}(g)$

Table I: Construction of $m_1(.)$ and $m_2(.)$.

where $f_{s,t}(g) \triangleq 1/(1+e^{-s(g-t)})$ is the sigmoid function; g is the criterion magnitude of the alternative under consideration; t is the abscissa of the inflection point of the sigmoid. The abscisses of inflection points are given by $t_c = g_j(b_h) - \frac{1}{2}(p_j(b_h) + q_j(b_h))$ and $t_{\bar{c}} = g_j(b_h) - \frac{1}{2}(p_j(b_h) + v_j(b_h))$ and the parameters s_c and $s_{\bar{c}}$ are given by[5] $s_c = 4/(p_j(b_h) - q_j(b_h))$ and $s_{\bar{c}} = 4/(v_j(b_h) - p_j(b_h))$.

From the setting of threshold parameters $p_j(b_h)$, $q_j(b_h)$ and $v_j(b_h)$ (the same as for ET method), it is easy to compute the parameters of the sigmoids (t_c, s_c) and $(t_{\bar{c}}, s_{\bar{c}})$, and thus to get the values of bba's $m_1(.)$ and $m_2(.)$ to compute $m^j_{ih}(.)$. We recommend to use the PCR5 fusion rule[6] since it offers a better management of conflicting bba's yielding to more specific results than with other rules. Based on this sigmoidal modeling, we get now from $m^j_{ih}(.)$ a fully consistent and efficient representation of local concordance $c_j(a_i, b_h)$, local discordance $d_j(a_i, b_h)$ and the local uncertainty $u_j(a_i, b_h)$ by considering:

$$\begin{cases} c_j(a_i, b_h) \triangleq m^j_{ih}(c) \in [0,1] \\ d_j(a_i, b_h) \triangleq m^j_{ih}(\bar{c}) \in [0,1] \\ u_j(a_i, b_h) \triangleq m^j_{ih}(c \cup \bar{c}) \in [0,1]. \end{cases} \tag{17}$$

Of course, a similar approach must be adapted (not reported here due to space limitation restraint) to

[3] Here we assume that Shafer's model holds, that is $c \cap \bar{c} = \emptyset$.

[4] with averaging rule, PCR5 rule, or Dempster-Shafer rule [14].

[5] The coefficient 4 appearing in s_c and $s_{\bar{c}}$ expressions comes from the fact that for a sigmoid of parameter s, the tangent at its inflection point is $s/4$.

[6] see [15] for details on PCR5 with many examples.

compute $c_j(b_h,a_i) = m^j_{hi}(c)$, $d_j(b_h,a_i) = m^j_{hi}(\bar{c})$ and $u_j(b_h,a_i) = m^j_{hi}(c \cup \bar{c})$.

Example 1: Let's consider only one alternative a_i and $g_j(.)$ in range $[0,100]$, and let's take $g_j(b_h) = 50$ and the following thresholds: $q_j(b_h) = 20$ (indifference threshold), $p_j(b_h) = 25$ (preference threshold) and $v_j(b_h) = 40$ (veto threshold) for the profile bound b_h. Then, the inflection points of the sigmoids $f_1(g) \triangleq f_{s_c,t_c}(g)$ and $f_2(g) \triangleq f_{-s_{\bar{c}},t_{\bar{c}}}(g)$ have the following abscisses: $t_c = 50-(25+20)/2 = 27.5$ and $t_{\bar{c}} = 50-(25+40)/2 = 17.5$ and parameters: $s_c = 4/(25-20) = 4/5 = 0.8$ and $s_{\bar{c}} = 4/(40-25) = 4/15 \approx 0.2666$. The construction of the consistent bba $m^j_{ih}(.)$ is obtained by the PCR5 fusion of the bba's $m_1(.)$ and $m_2(.)$ given in Table I. The result is shown in Fig. 2.

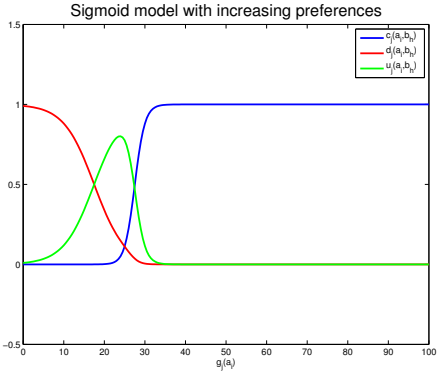


Figure 2: $m^j_{ih}(.)$ corresponding to partial indices.

The blue curve corresponds to $c_j(a_i,b_h)$, the red plot corresponds to $d_j(a_i,b_h)$ and the green plot to $u_j(a_i,b_h)$ when $g_j(a_i)$ varies in $[0;100]$. $c_j(b_h,a_i)$, $d_j(b_h,a_i)$ and $u_j(b_h,a_i)$ can easily be obtained by mirroring (horizontal flip) the curves around the vertical axis at the mid-range value $g_j(a_i) = 50$.

B. SET-Step 2: Global indices

As explained in SET-Step 1, the partial indices are encapsulated in bba's $m^j_{ih}(.)$ for alternative a_i versus profile b_h ($a_i$vs.b_h), and encapsulated in bba's $m^j_{hi}(.)$ for profile b_h versus alternative a_i ($b_h$vs.a_i). In SET, the global indices $c(a_i,b_h)$, $d(a_i,b_h)$ and $u(a_i,b_h)$ are obtained by the fusion of the n_g bba's $m^j_{ih}(.)$. Similarly, the global indices $c(b_h,a_i)$, $d(b_h,a_i)$ and $u(b_h,a_i)$ are obtained by the fusion of the n_g bba's $m^j_{hi}(.)$. More precisely, one must compute:

$$\begin{cases} m_{ih}(.) = [m^1_{ih} \oplus m^2_{ih} \oplus \ldots \oplus m^{n_g}_{ih}](.) \\ m_{hi}(.) = [m^1_{hi} \oplus m^2_{hi} \oplus \ldots \oplus m^{n_g}_{hi}](.) \end{cases} \quad (18)$$

To take into account the weighting factor w_j of the criterion valued by $g_j(.)$, we suggest to use as fusion operator $\oplus$ either:

- the weighting averaging fusion rule (as in ET method) which is simple and compatible with probability calculus and Bayesian reasoning,
- or the more sophisticated operator defined by the PCR5 fusion rule adapted for importance discounting presented in details in [16] which belongs to the family of non-Bayesian fusion operators.

Once the bba's $m_{ih}(.)$ and $m_{hi}(.)$ have been computed, the global indices are defined by:

$$\begin{cases} c(a_i,b_h) \triangleq m_{ih}(c)\alpha(a_i,b_h) \\ d(a_i,b_h) \triangleq m_{ih}(\bar{c})\beta(a_i,b_h) \\ u(a_i,b_h) \triangleq 1 - c(a_i,b_h) - d(a_i,b_h). \end{cases} \quad (19)$$

The discounting factors $\alpha(a_i,b_h)$ and $\beta(a_i,b_h)$ are defined by

$$\alpha(a_i,b_h) \triangleq \begin{cases} 1 & \text{if } \mathbf{V}_\alpha = \emptyset \\ \prod_{j\in\mathbf{V}_\alpha} \frac{1-d_j(a_i,b_h)}{1-m_{ih}(c)} & \text{if } \mathbf{V}_\alpha \neq \emptyset \end{cases} \quad (20)$$

$$\beta(a_i,b_h) \triangleq \begin{cases} 1 & \text{if } \mathbf{V}_\beta = \emptyset \\ \prod_{j\in\mathbf{V}_\beta} \frac{1-c_j(a_i,b_h)}{1-m_{ih}(\bar{c})} & \text{if } \mathbf{V}_\beta \neq \emptyset \end{cases} \quad (21)$$

$$\text{with} \quad \begin{cases} \mathbf{V}_\alpha \triangleq \{j \in \mathbf{J} | d_j(a_i,b_h) > m_{ih}(c)\} \\ \mathbf{V}_\beta \triangleq \{j \in \mathbf{J} | c_j(b_h,a_i) > m_{ih}(\bar{c})\} \end{cases} \quad (22)$$

$c(b_h,a_i)$, $d(b_h,a_i)$ and $u(b_h,a_i)$ are similarly computed using dual formulas of (19)–(22).

The belief and plausibility of the outranking propositions $X =$ "$a_i > b_h$" and $Y =$ "$b_h > a_i$" are then given by

$$\begin{cases} \text{Bel}(X) = c(a_i,b_h) \\ \text{Bel}(Y) = c(b_h,a_i) \end{cases} \quad (23)$$

$$\text{and} \quad \begin{cases} \text{Pl}(X) = 1 - d(a_i,b_h) = c(a_i,b_h) + u(a_i,b_h) \\ \text{Pl}(Y) = 1 - d(b_h,a_i) = c(b_h,a_i) + u(b_h,a_i) \end{cases} \quad (24)$$

C. SET-Step 3: Probabilized outranking

We have seen in SET-Step 2 that the outrankings $X =$ "$a_i > b_h$" and $Y =$ "$b_h > a_i$" can be characterized by their imprecise probabilities $P(X) \in [\text{Bel}(X); \text{Pl}(X)]$ and $P(Y) \in [\text{Bel}(Y); \text{Pl}(Y)]$. Figure 3 shows an example with $P(X) \in [0.2; 0.8]$ and $P(Y) \in [0.1; 0.5]$

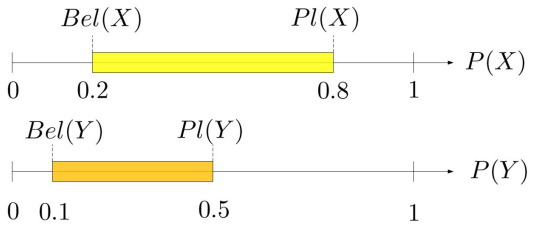


Figure 3: Imprecise probabilities of outrankings.

Solving the outranking problem consists in choosing (deciding) if finally X dominates Y (in such case we must decide X as being the valid outranking), or if Y dominates X (in such case we decide Y as being the valid outranking). Unfortunately, such hard (binary) assignment cannot be done in general[7] because it must be drawn from the unknown probabilities $P(X)$ in $[\text{Bel}(X); \text{Pl}(X)]$ and $P(Y)$

[7]but in cases where the bounds of probabilities $P(X)$ and $P(Y)$ do not overlap.

in $[\text{Bel}(Y);\text{Pl}(Y)]$ where a partial overlapping is possible between intervals $[\text{Bel}(X);\text{Pl}(X)]$ and $[\text{Bel}(Y);\text{Pl}(Y)]$ (see Fig. 3). A soft (probabilized) outranking solution is possible by computing the probability that X dominates Y (or that Y dominates X) by assuming uniform distribution of unknown probabilities between their lower and upper bounds. To get the probabilized outrankings, we just need to compute $P_{X>Y} \triangleq P(P(X) > P(Y))$ and $P_{Y>X} \triangleq P(P(Y) > P(X))$ which are precisely computable by the ratio of two polygonal areas, or can be estimated using sampling techniques.

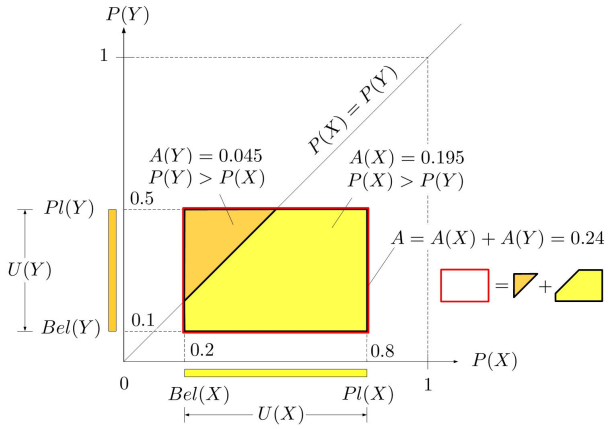


Figure 4: Probabilization of outranking.

$$\text{More precisely } \begin{cases} P_{X>Y} = A(X)/(A(X)+A(Y)) \\ P_{Y>X} = A(Y)/(A(X)+A(Y)) \end{cases} \quad (25)$$

where $A(X)$ is the partial area of the rectangle $A = U(X) \times U(Y)$ under the line $P(X) = P(Y)$ (yellow area in Fig. 4) and $A(Y)$ is the area of the rectangle $A = U(X) \times U(Y)$ above the line $P(X) = P(Y)$ (orange area in Fig. 4). Of course, $A = A(X) + A(Y)$ and $P_{X>Y} = 1 - P_{Y>X}$. As a final result for the example of Fig. 3, and according to (25) and Fig. 4, we finally get the following probabilized outrankings:

$$\begin{cases} a_i > b_h \text{ with probability} P_{X>Y} = 0.195/0.24 = 0.8125 \\ b_h > a_i \text{ with probability} P_{Y>X} = 0.045/0.24 = 0.1825 \end{cases}$$

For notation convenience, we denote the probabilities of outrankings as $P_{ih} \triangleq P_{X>Y}$ with $X =$ "$a_i > b_h$" and $Y =$ "$b_h > a_i$". Reciprocally, we denote $P_{hi} \triangleq P_{Y>X} = 1 - P_{ih}$.

D. SET-Step 4: Soft assignment procedure

From the probabilized outrankings obtained in SET-Step 3, we are now able to make directly the soft assignment of alternatives a_i to categories C_h defined by their profiles b_h. This is easily obtained by the combinatorics of all possible sequences of outrankings taking into account their probabilities. Moreover, this soft assignment mechanism provides also the probability $\delta_i \triangleq P(a_i \to \emptyset)$ reflecting the impossibility to make a coherent outranking. Our soft assignment procedure doesn't require arbitrary choice of attitude contrariwise to what is proposed in the classical ET method. For simplicity, we present the soft assignment procedure in the example 2 below, which can be adapted to any number $n_h \geq 2$ of categories.

Example 2: Let's consider one alternative a_i to be assigned to categories C_1, C_2 and C_3 based on multiple criteria (taking into account indifference, preference and veto conditions) and intermediate profiles b_1 and b_2. Because b_0 and b_3 are the min and max profiles, one has always $P(X_{i0} =$ "$a_i > b_0$"$) = 1$ and $P(X_{i3} =$ "$a_i > b_3$"$) = 0$. Let's assume that at the SET-Step 3 one gets the following soft outranking probabilities P_{ih} as given in Table II.

Profiles $b_h \to$ **Outranking probas** $\downarrow$	b_0	b_1	b_2	b_3
P_{ih}	1	0.7	0.2	0

Table II: Soft outranking probabilities.

From combinatorics, only the following outranking sequences $S_k(a_i)$, $k = 1,2,3,4$ can occur with non null probabilities $P(S_k(a_i))$ as listed in Table III, where $P(S_k(a_i))$

Profiles $b_h \to$ **Outrank sequences** $\downarrow$	b_0	b_1	b_2	b_3	$P(S_k(a_i))$ $\downarrow$
$S_1(a_i)$	$>$	$>$	$>$	$<$	0.14
$S_2(a_i)$	$>$	$>$	$<$	$<$	0.56
$S_3(a_i)$	$>$	$<$	$<$	$<$	0.24
$S_4(a_i)$	$>$	$<$	$>$	$<$	0.06

Table III: Probabilities of outranking sequences.

have been computed by the product of the probability of each outranking involved in the sequence, that is:

$$P(S_1(a_i)) = 1 \times 0.7 \times 0.2 \times 1 = 0.14$$
$$P(S_2(a_i)) = 1 \times 0.7 \times (1-0.2) \times 1 = 0.56$$
$$P(S_3(a_i)) = 1 \times (1-0.7) \times (1-0.2) \times 1 = 0.24$$
$$P(S_4(a_i)) = 1 \times (1-0.7) \times 0.2 \times 1 = 0.06$$

The assignment of a_i into a category C_h delimited by bounds b_{h-1} and b_h depends on the occurrence of the outranking sequences. Given $S_1(a_i)$ with probability $P(S_1(a_i)) = 0.14$, a_i must be assigned to C_3 because a_i outranks b_0, b_1 and b_2; Given $S_2(a_i)$ with probability 0.56, a_i must be assigned to C_2 because a_i outranks only b_0 and b_1; Given $S_3(a_i)$ with probability 0.24, a_i must be assigned to C_1 because a_i outranks only b_0. Given $S_4(a_i)$ with probability 0.06, a_i cannot be reasonably assigned to categories because of inherent inconsistency of the outranking sequence $S_4(a_i)$ since a_i cannot outperform b_2 and simultaneously underperform b_1 because by profile ordering one has $b_2 > b_1$. Therefore the inconsistency indicator is given by $\delta_i = P(a_i \to \emptyset) = P(S_4(a_i)) = 0.06$. Finally, the soft assignment probabilities $P(a_i \to C_h)$ and the inconsistency indicator obtained by SET-Step 4 are given in Table IV.

Categories $C_h \rightarrow$ **Assignment probas** $a_i \downarrow$	C_1	C_2	C_3	$\emptyset$
$P(a_i \rightarrow C_h)$	0.24	0.56	0.14	$\delta_i = 0.06$

Table IV: SET Soft Assignment result.

IV. Application example : environmental context

In this section, we compare ET and SET methods applied to an assignment problem related to an environmental context proposed originally in [8]. It corresponds to the choice of the location of an urban waste resource recovery disposal which aims to re-use the recyclable part of urban waste produced by several communities. Indeed, this disposal must collect at least $20000m^3$ of urban waste per year to be economically viable. It must be a collective unit and the best possible location has to be identified. Each community will have to bring its urban waste production to the disposal: the transport costs are valuated in tons by kilometer per year ($t.km/year$). Building such a disposal is generally not easily accepted by population, particularly when the environmental inconveniences are already high. This initial environmental status is measured by a specific criterion. Building an urban waste disposal implies to use a wide area that could be used for other activities such as a sport terrain, touristic equipments, a natural zone, etc. This competition with other activities is measured by a specific criterion.

A. Alternatives, criteria and profiles definition

In our example, 7 possible locations (alternatives/choices) a_i, $i = 1, 2, \ldots, 7$, for urban waste resource recovery disposal are compared according to the following 5 criteria g_j, $j = 1, 2, \ldots, 5$:

$g_1 =$ Terrain price (decreasing preference);

$g_2 =$ Transport costs (decreasing preference);

$g_3 =$ Environment status (increasing preference);

$g_4 =$ Impacted population (increasing preference);

$g_5 =$ Competition activities (increasing preference).

- Price of terrain (g_1) is expressed in $€/m^2$ with decreasing preferences (the lower is the price, the higher is the preference);
- Transport costs (g_2) are expressed in $t.km/year$ with decreasing preferences (the lower is the cost, the higher is the preference);
- The environment status (g_3) corresponds to the initial environmental inconvenience level expressed by population with an increasing direction of preferences. The higher is the environment status, the lower are the initial environmental inconveniences. It is rated with an integer between 0 and 10 (highest environment status corresponding to the lowest initial environmental inconveniences);
- Impacted population (g_4) is an integrated criterion to measure negative effects based on subjective and qualitative criteria. It corresponds to the status of the environment with an increasing direction of preferences. The higher is the evaluation, the lower are the negative effects. It is rated with an real number between 0 (great number of impacted people) and 10 (very few people impacted);
- Activities competition (g_5) is an integrated criterion, evaluated by a real number, that measures the competition level between activities with an increasing direction of preferences. The higher is the evaluation, the lower is the competition with other activities on the planned location (tourism, sport, natural environment ...).

The evaluations of the 7 alternatives are summarized in Table V, and he alternatives (possible locations) are compared to the 2 decision profiles b_1 and b_2 described in Table VI. The weights, indifference, preference and veto thresholds for criteria g_j are described in Table VII.

Criteria $g_j \rightarrow$ **Choices** $a_i \downarrow$	g_1 ($€/m^2$)	g_2 ($t \cdot km$/year)
a_1	-120	-284
a_2	-150	-269
a_3	-100	-413
a_4	-60	-596
a_5	-30	-1321
a_6	-80	-734
a_7	-45	-982

(a) Choices a_i and criteria g_1 and g_2.

Criteria $g_j \rightarrow$ **Choices** $a_i \downarrow$	g_3 $\{0, 1, \ldots, 10\}$	g_4 $[0, 10]$	g_5 $\{0, 1, \ldots, 100\}$
a_1	5	3.5	18
a_2	2	4.5	24
a_3	4	5.5	17
a_4	6	8.0	20
a_5	8	7.5	16
a_6	5	4.0	21
a_7	7	8.5	13

(b) Choices a_i and criteria g_3, g_4 and g_5.

Table V: Inputs of ET (7 alternatives according to 5 criteria).

Profiles $b_h \rightarrow$ **Criteria** $g_j \downarrow$	b_1	b_2
g_1: $€/m^2$	-100	-50
g_2: $t \cdot km$/year	-1000	-500
g_3: $\{0, 1, \ldots, 10\}$	4	7
g_4: $[0, 10]$	4	7
g_5: $\{0, 1, \ldots, 100\}$	15	20

Table VI: Evaluation profiles.

Thresholds $\rightarrow$ **Criteria** $g_j \downarrow$	w_j (weight)	q_j (indifference)	p_j (preference)	v_j (veto)
g_1: $€/m^2$	0.25	15	40	100
g_2: $t \cdot km$/year	0.45	80	350	850
g_3: $\{0, 1, \ldots, 10\}$	0.10	1	3	5
g_4: $[0, 10]$	0.12	0.5	3.5	4.5
g_5: $\{0, 1, \ldots, 100\}$	0.08	1	5	8

Table VII: Thresholds.

B. Results of classical ELECTRE TRI

After applying ET-Steps 1 and 3 of the classical ET method described in Section II with a $\lambda = 0.75$ for the λ-cut strategy, one gets the outranking relations listed in Table VIII.

The final hard assignments obtained by ET method using the pessimistic and optimistic attitudes are listed in Table IX.

C. Results of the new Soft ELECTRE TRI

After applying SET-Steps 1 and 3 of the SET method[8] described in Section III, one gets the probabilities of soft outrankings listed in Table X.

[8] We have used here the PCR5 fusion rule with importance discounting [16], and a sampling technique to compute the probabilities P_{ih}.

	b_0	b_1	b_2	b_3
a_1	$>$	$>$	$<$	$<$
a_2	$>$	R	R	$<$
a_3	$>$	$>$	$<$	$<$
a_4	$>$	$>$	I	$<$
a_5	$>$	R	$<$	$<$
a_6	$>$	$>$	$<$	$<$
a_7	$>$	$>$	$<$	$<$

Table VIII: Outranking relations obtained with ET ($\lambda = 0.75$).

	C_1	C_2	C_3
a_1	0	1	0
a_2	1	0	0
a_3	0	1	0
a_4	0	1	0
a_5	1	0	0
a_6	0	1	0
a_7	0	1	0

(a) Pessimistic attitude.

	C_1	C_2	C_3
a_1	0	1	0
a_2	0	0	1
a_3	0	1	0
a_4	0	0	1
a_5	0	1	0
a_6	0	1	0
a_7	0	1	0

(b) Optimistic attitude.

Table IX: Hard assignments obtained with ET ($\lambda = 0.75$).

Profiles $b_h \rightarrow$ **Outranking probas** $\downarrow$	b_0	b_1	b_2	b_3
P_{1h}	1	0.9858	0.6211	0
P_{2h}	1	0.8908	0.1812	0
P_{3h}	1	0.9999	0.0570	0
P_{4h}	1	1.0000	0.0807	0
P_{5h}	1	0.2142	0.0145	0
P_{6h}	1	0.9996	0.0006	0
P_{7h}	1	0.9975	0.0106	0

Table X: Probabilities of soft outranking relations by SET.

The final soft assignments obtained by the SET method are listed in Table XI.

	C_1	C_2	C_3	$\emptyset$
a_1	0.0054	0.3735	0.6123	$\delta_1 = 0.0088$
a_2	0.0894	0.7294	0.1614	$\delta_2 = 0.0198$
a_3	0.0001	0.9429	0.0570	$\delta_3 = 0$
a_4	0	0.9193	0.0807	$\delta_4 = 0$
a_5	0.7744	0.2111	0.0031	$\delta_5 = 0.0114$
a_6	0.0004	0.9990	0.0006	$\delta_6 = 0$
a_7	0.0025	0.9869	0.0106	$\delta_7 = 0$

Table XI: SET Soft Assignment matrix $[P(a_i \rightarrow C_h)]$.

D. Discussion

From Table XI, we can get a hard assignment solution (if needed) by assigning each alternative to the category corresponding to the maximum of $P(a_i \rightarrow C_h)$, $h = 1, 2, \ldots$. With SET, it is also theoretically possible to "assign" a_i to none category if δ_i (inconsistency level) is too high. The soft assignments for a_i, $i = 3, \ldots 7$ (see Tables IX,XI) are compatible with the hard assignments with the pessimistic or the optimistic attitudes. In fact, only the soft assignments for a_1 and a_2 having the highest probabilities $P(a_1 \rightarrow C_3) = 0.6123$ and $P(a_2 \rightarrow C_2) = 0.7294$ appear incompatible with ET hard assignments (pessimistic or optimistic). The discrepancy between these soft and hard assignments solutions is not due to SET method but comes from the arbitrary choice of the level of the λ-cut strategy used in ET method. Another arbitrary choice of λ-cut will generate different ET hard assignments which can in fact become fully compatible with SET soft assignments. For example, if one takes $\lambda = 0.5$, it can be verified that SET soft assignments are now compatible with ET hard assignments for all alternatives in this example. The soft assignments approach of SET is interesting since it doesn't depend on λ values even if the influence of both sigmoids parameters definition, choice of fusion rule, probabilisation method ... could be further studied.

V. Conclusions

A new outranking sorting method, called Soft ELECTRE TRI (SET), inspired from the classical ELECTRE TRI and based on beliefs functions and advanced fusion techniques is proposed. SET method uses the same inputs as ET (same criteria and thresholds definitions) but in a more effective way and provides a soft (probabilized) assignment solution. SET eliminates the inherent problem of classical ET due to the arbitrary choice of a λ-cut strategy which forces to adopt either a pessimistic or optimistic attitude for the final hard assignment of alternatives to categories. The interest of SET over ET method is demonstrated on a preexisting environmental context scenario.

References

[1] R. Keeney, H. Raiffa, *Decisions with multiple objectives:preferences and values trade-offs*, John Wiley and Sons, New York, 1976.

[2] J. Dyer, *MAUT - Multiattribute Utility theory*, in [17], pp. 263–295.

[3] B. Roy, *Main sources of inaccurate determination, uncertainty and imprecision in decision models*, Math. and Comput. Modelling, Vol. 12, No. 10-11, pp. 1245–1254, 1989.

[4] W. Yu, *Aide multicritère a la décision dans le cadre de la problématique du tri: Concepts, méthodes et applications*, Ph.D Thesis, University Paris-Dauphine, Paris, France, 1992.

[5] B. Roy, *Classement et choix en présence de points de vue multiples (la méthode ELECTRE)*, RIRO, Vol. 8, pp. 57-75, 1968.

[6] J. Figueira, V. Mousseau, B. Roy, *ELECTRE methods*, in [17], pp. 133–162.

[7] LAMSADE web site: http://www.lamsade.dauphine.fr/mcda/biblio/

[8] L.Y. Maystre, J. Pictet, J.M. Simos, *Méthodes multicritères ELECTRE*, Presses polytechniques et universitaires romandes, Lausanne, 1994.

[9] V. Mousseau, R. Slowinski, P. Zielniewicz, *ELECTRE TRI 2.0a user's manual*, Cahier et documents du Lamsade, Paris-Dauphine Univ., 1999.

[10] T. Tervonen, et al., *A stochastic method for robustness analysis in sorting problems*, European Journal of Operational Research, 192, pp. 236–242, 2009.

[11] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[12] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vol. 1-3, American Research Press, 2004–2009. http://www.gallup.unm.edu/~smarandache/DSmT.htm

[13] J. Dezert, J.M. Tacnet, *Sigmoidal Model for Belief Function-based ELECTRE TRI Method*, Proc. of Belief 2012, France, May 2012.

[14] J. Dezert, F. Smarandache, *An introduction to DSmT*, in [12], Vol. 3, pp. 3–73.

[15] J. Dezert, F. Smarandache, *Proportional Conflict Redistribution Rules for Information Fusion*, in [12], Vol. 2, pp. 3–68.

[16] F. Smarandache, J. Dezert, J.M.Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, in Proc. of Fusion 2010 Conf., Seattle, USA, July 2010.

[17] J. Figueira, S. Greco, and M. Ehrgott (Editors), *Multiple Criteria Decision Analysis: State of the Art Surveys*, Springer Verlag, 2005.

On The Validity of Dempster-Shafer Theory

Jean Dezert
Pei Wang
Albena Tchamova

Originally published as Dezert J., Wang P., Tchamova A., *On The Validity of Dempster-Shafer Theory,* in Proc. of Fusion 2012, Singapore, July 2012, and reprinted with permission.

***Abstract*—We challenge the validity of Dempster-Shafer Theory by using an emblematic example to show that DS rule produces counter-intuitive result. Further analysis reveals that the result comes from a understanding of evidence pooling which goes against the common expectation of this process. Although DS theory has attracted some interest of the scientific community working in information fusion and artificial intelligence, its validity to solve practical problems is problematic, because it is not applicable to evidences combination in general, but only to a certain type situations which still need to be clearly identified.**

Keywords: Dempster-Shafer Theory, DST, Mathematical Theory of Evidence, belief functions.

I. Introduction

Dempster-Shafer Theory (DST), also known as the Theory of Evidence or the Theory of Belief Functions, was introduced by Shafer in 1976 [1], based on Dempster's previous works [2]–[4]. This theory offers an elegant theoretical framework for modeling uncertainty, and provides a method for combining distinct bodies of evidence collected from different sources. In the past more than three decades, DST has been used in many applications, in fields including information fusion, pattern recognition, and decision making [5].

Even so, starting from Zadeh's criticism [6]–[8], many questions have arisen about the validity and the consistency of DST when combining uncertain and conflicting evidences expressed as basic belief assignments (bba's). Beside Zadeh's example, there have been several detailed analysis on this topic by Lemmer [9], Voorbraak [10] and Wang [11]. Other authors like Pearl [12], [13] and Walley [14], and more recently Gelman [15], have also warned the "belief function community" about the validity of Dempster-Shafer's rule (DS rule for short) for combining distinct pieces of evidences based on different analyses and contexts. Since the mid-1990's, many researchers and engineers working with belief functions in applications have observed and recognized that DS rule is problematic for evidence combination, specially when the sources of evidence are high conflicting.

In response to this challenge, various attempts have been made to circumvent the counter-intuitive behavior of DS rule. They either replace Dempster-Shafer's rule by alternative rules, listed for example in [16] (Vol. 1), or apply novel semantic interpretations to the functions [16]–[18].

Before going further in our discussion, let us recall two of Shafer's statements about DST:

> The burden of our theory is that this rule [Dempster's rule of combination] corresponds to the pooling of evidence: if the belief functions being combined are based on entirely distinct bodies of evidence and the set Θ discerns the relevant interaction between those bodies of evidence, then the orthogonal sum gives degree of belief that are appropriate on the basis of combined evidence. [1] (p. 6)
>
> This formalism [whereby propositions are represented as subsets of a given set] is most easily introduced in the case where we are concerned with the true value of some quantity. If we denote the quantity by θ and the set of its possible values by Θ, then the propositions of interest are precisely those of the form "The true value of θ is in T," where T is a subset of Θ. [1] (p. 36)

These two statements are very important since they are related to two fundamental questions on DST that are central in this discussion on the validity of DS theory:

1) What is the meaning of "pooling of evidence" used by Shafer? Does it correspond to an experimental protocol?
2) When "the true value of θ is in T" is asserted by a source of evidence, are we getting absolute truth (based on the whole knowledge accessible by everyone eventually) or relative truth (based on the partial knowledge accessible by the source at the moment)?

This paper starts with a very emblematic example to show what we consider as really problematic in DS rule behavior, which corresponds to the possible "dictatorial power" of a source of evidence with respect to all others and thus reflecting the minority opinion. We demonstrate that the problem is in fact not merely due to the level of conflict between sources to combine, but comes from the underlying interpretations of *evidence* and *degree of belief* on which the combination rule is based. Such interpretations do not agree with the common usage of those notions where an opinion based on certain evidence can be revised by (informative) evidence from other sources.

This work is based on our preliminary ideas presented in the Spring School on Belief Functions Theory and Applications (BFTA) in April 2011 [19], and on many fruitful discussions with colleagues using belief functions. Their stimulating comments, especially when they disagree, help us to clarify and present our ideas.[1] In Section II we briefly recall basics of DST and DS rule. In Section III, we describe the example and its strange (counter-intuitive) result. In Section IV we present a general analysis on the validity of DST, and we conclude our analysis in Section V.

II. Basics of DST

Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ be a frame of discernment of a problem under consideration containing n distinct elements θ_i, $i = 1, \ldots, n$.

A basic belief assignment (*bba*, also called a belief mass function) $m(.) : 2^\Theta \rightarrow [0, 1]$ is a mapping from the power set of Θ (i.e. the set of subsets of Θ), denoted 2^Θ, to $[0, 1]$, that must satisfy the following conditions: 1) $m(\emptyset) = 0$, i.e. the mass of empty set (impossible event) is zero; 2) $\sum_{X \in 2^\Theta} m(X) = 1$, i.e. the mass of belief is normalized to one. Here $m(X)$ represents the mass of belief exactly committed to X. An element $X \in 2^\Theta$ is called a focal element if and only if $m(X) > 0$. The set $\mathcal{F}(m) \triangleq \{X \in 2^\Theta | m(X) > 0\}$ of all focal elements of a bba $m(.)$ is called the core of the bba. By definition, a Bayesian bba $m(.)$ is a bba having only focal elements of cardinality 1. The vacuous bba characterizing full ignorance is defined by $m_v(.) : 2^\Theta \rightarrow [0; 1]$ such that $m_v(X) = 0$ if $X \neq \Theta$, and $m_v(\Theta) = 1$.

From any bba $m(.)$, the belief function $Bel(.)$ and the plausibility function $Pl(.)$ are defined as $\forall X \in 2^\Theta : Bel(X) = \sum_{Y|Y \subseteq X} m(Y)$ and $Pl(X) = \sum_{Y|X \cap Y \neq \emptyset} m(Y)$. $Bel(X)$ represents the whole mass of belief that comes from all subsets of Θ included in X. $Pl(X)$ represents the whole mass of belief that comes from all subsets of Θ compatible with X (i.e., those intersecting X).

The DS rule of combination [1] is an operation denoted $\oplus$, which corresponds to the normalized conjunction of mass functions. Based on Shafer's description, given two independent and distinct sources of evidences characterized by bba $m_1(.)$ and $m_2(.)$ on the same frame of discernment Θ, their combination is defined by $m_{DS}(\emptyset) = 0$, and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{DS}(X) = [m_1 \oplus m_2](X) = \frac{m_{12}(X)}{1 - K_{12}} \tag{1}$$

where

$$m_{12}(X) \triangleq \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) \tag{2}$$

corresponds to the conjunctive consensus on X between the two sources of evidence. K_{12} is the total *degree of conflict* between the two sources of evidence defined by

$$K_{12} \triangleq m_{12}(\emptyset) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2) \tag{3}$$

When $K_{12} = m_{12}(\emptyset) = 1$, the two sources are said in total conflict and their combination cannot be applied since DS rule (1) is mathematically undefined, because of 0/0 indeterminacy [1]. DS rule is commutative and associative, which makes it attractive from engineering implementation standpoint, since the combinations of sources can be done sequentially instead globally and the order doesn't matter. Moreover, the vacuous bba is a neutral element for the DS rule, i.e. $[m \oplus m_v](.) = [m_v \oplus m](.) = m(.)$ for any bba $m(.)$ defined on 2^Θ, which seems to be an expected[2] property, i.e. a full ignorant source doesn't impact the fusion result.

The conditioning of a given bba $m(.)$ by a conditional element $Z \in 2^\Theta \setminus \{\emptyset\}$ has been also proposed by Shafer [1]. This function $m(.|Z)$ is obtained by DS combination of $m(.)$ with the bba $m_Z(.)$ only focused on Z, i.e. such that $m_Z(Z) = 1$. For any element X of the power set 2^Θ this is mathematically expressed by

$$m(X|Z) = [m \oplus m_Z](X) = [m_Z \oplus m](X) \tag{4}$$

It has been proved [1] (p. 67) that this rule of conditioning expressed in terms of plausibility functions yields to the formula

$$Pl(X|Z) = Pl(X \cap Z)/Pl(Z) \tag{5}$$

which is very similar to the well-known Bayes formula $P(X|Z) = P(X \cap Z)/P(Z)$. Partially because of this, DST has been widely considered as a generalization of Bayesian inference [3], [4], or equivalently, that probability theory is a special case of the Mathematical Theory of Evidence when manipulating Bayesian bba's.

Despite of the appealing properties of DS rule, its apparent similarity with Bayes formula for conditioning, and many attempts to justify its foundations, several challenges on the theory's validity have been put forth in the last decades, and remain unanswered. For instance, an experimental protocol to test DST was proposed by Lemmer in 1985 [9], and his analysis shows an inherent paradox (contradiction) of DST. Following a different approach, an inconsistency in the fundamental postulates of DST was proved by Wang in 1994 [11]. Some other related works questioning the validity of DST based on different argumentations have been listed in the introduction of this paper.

In the following, we identify the origin of the problem of DS rule under the common interpretation of the "pooling" of evidence, and why it is very risky to use it in very sensitive applications, specially where security, defense and safety are involved.

[1] Our presentation is not based on a previous statistical argumentation developed in [20], since it appears for some strong proponents of DST as an invalid approach to criticize DS rule. In this paper we adopt a simpler argumentation based only on common sense and simple considerations manipulating witnesses reports.

[2] A detailed discussion about this "expected" property can be found in [20].

III. A SIMPLE EXAMPLE AND ITS STRANGE RESULT

To see the problem in combining evidence with DS rule, let us analyze an emblematic example. Consider a frame of discernment with three elements only, $\Theta = \{A, B, C\}$, satisfying Shafer's request, i.e. the elements of the frame are truly exhaustive and exclusive. As in Zadeh's example, we interpret the problem as medical diagnosis, where A, B and C correspond to three distinct pathologies (say $A =$ brain tumor, $B =$ concussion and $C =$ meningitis) of a patient. In such a situation, it is reasonable to assume that these pathologies do not occur simultaneously, so Shafer's assumptions truly hold.

We suppose that two distinct doctors (or more generally, two witnesses) provide their own medical diagnostic (or more generally, a testimony) of the same patient, based on their own knowledges and expertises, after analyzing symptoms, IRM images, or any useful medical results. The diagnostics (testimonies) of the two distinct sources of evidences correspond to the two non-Bayesian bba's given by the doctors listed in Table I. The parameters a, b_1, and b_2 can take any value, as long as $a \in [0, 1]$, $b_1, b_2 > 0$, and $b_1 + b_2 \in [0, 1]$.

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	b_1
C	0	$1-b_1-b_2$
$A \cup B \cup C$	0	b_2

Table I
INPUT BBA'S $m_1(.)$ AND $m_2(.)$.

The two distinct sources are assumed to be truly independent (the diagnostic of Doctor 1 is done independently of the diagnostic of Doctor 2 and from different medical results, images supports, etc, and conversely) so that we are allowed to apply DS rule to combine the two bba's $m_1(.)$ and $m_2(.)$. Both doctors are also assumed to have the same level of expertise and they are equally reliable. Note that in this very simple parametric example the focal elements of bba's are not nested (consonant), and there really does exist a conflict between the two sources (as it will be shown in the derivations). It is worth to note also that the two distinct sources are truly informative since none of them corresponds to the vacuous belief assignment representing a full ignorant source, so it is reasonable to expect for both bba's to be taken into account (and to have an impact) in the fusion process. Here we use the notion of "conflict" as defined by Shafer in [1] (p. 65) and recalled by (3).

When applying DS rule of combination, one gets:

1) Using the conjunctive operator:

$$m_{12}(A) = a(b_1 + b_2) \tag{6}$$

$$m_{12}(A \cup B) = (1-a)(b_1 + b_2) \tag{7}$$

$$K_{12} = m_{12}(\emptyset) = 1 - b_1 - b_2 \qquad \text{(conflicting mass)} \tag{8}$$

2) and After normalizing by $1 - K_{12} = b_1 + b_2$, the final result is as follows:

$$\begin{aligned} m_{DS}(A) &= \frac{m_{12}(A)}{1-K_{12}} = \frac{a(b_1+b_2)}{b_1+b_2} \\ &= a = m_1(A) \end{aligned} \tag{9}$$

$$\begin{aligned} m_{DS}(A \cup B) &= \frac{m_{12}(A \cup B)}{1-K_{12}} = \frac{(1-a)(b_1+b_2)}{b_1+b_2} \\ &= 1 - a = m_1(A \cup B) \end{aligned} \tag{10}$$

Surprisingly, after combining the two sources of evidences with Dempster-Shafer's rule, we see that in this case the medical diagnostic of Doctor 2 doesn't count at all, because one gets $m_{DS}(.) = m_1(.)$. Though Doctor 2 is not a fully ignorant source and he/she has same reliability as Doctor 1, nevertheless his/her report (whatever it is when changing values of b_1 and b_2) doesn't count. We see that the level of conflict $K_{12} = 1-b_1-b_2$ between the two medical diagnostics doesn't matter in fact in the DS fusion process, since it can be chosen at any high or low level, depending on the choice of $b_1 + b_2$. Based on DST analysis, the Doctor 2 plays the same role as a vacuous/ignorant source of evidence even if he/she is informative (not vacuous), and truly conflicting (according to Shafer's definition) with Doctor 1.

This result goes against common sense. It casts serious doubt on the validity of DS rule, as well as its usefulness for applications, and interrogates on the real meaning of Shafer's pooling of evidence process. This example seems more crucial than the examples discussed in the existing literature in showing intolerable flaws in DST behavior, since in this example the level of conflict (whatever it is) between the sources doesn't play a role at all, so that it cannot be argued that in such a case DS must not be applied because of the high conflicting situation. In fact such a situation can occur in real applications and is not anecdotal, and the results obtained by DS rule can yield dramatical consequences. From Zadeh's example [6] and all the debates about it in the literature, it has been widely (though not completely) admitted that DS is not recommended when the conflict between sources is high. Our example brings out a more important question since it reveals that the problem of the behavior of DS rule is not due to the (high) level of conflict between the sources, but from something else — we can choose a low conflict level, but the result is still the same, so the problem remains.

We can see the situation better by generalizing from this example. What make this example special and emblematic of DS behavior is the fact that $Pl_1(C) = 0$. It not only means that Doctor 1 completely rules out the possibility of C, but also that this opinion cannot be changed by taking new evidence into consideration. This is the case, because according to Shafer's definition [1] (p. 43), $Pl_1(C) = 0$ means for every $X \in 2^\Theta$ that $X \cap C \neq \emptyset$, $m_1(X) = 0$. When DS rule is applied to combine $m_1(.)$ and an arbitrary $m'(.)$, for every $Y \in 2^\Theta$ that $Y \cap C \neq \emptyset$, $m_{DS}(Y) = 0$, because it is the sum of some products, each of them take one of the above $m_1(X)$ as a factor. Consequently, $Pl_{DS}(C) = 0$, no matter what the other body of evidence is. Actually in such situations DS rule

doesn't perform a fusion between sources' opinions, but an exclusion, ruling out the conflicting hypothesis considered by the second source.

Put it in another way, the effective frame of discernment of Doctor 1 is not really $\{A, B, C\}$, but $\{A, B\}$, because the pathology C has been ruled out of the frame by Doctor 1, since the focal elements of $m_1(.)$ are A and $A \cup B$ only. The above analysis tells us that when different supports (i.e. sets of focal elements) are combined according to DS rule, the resulting bba will be defined in the intersection of the supports of each source, under the condition that it is not empty (otherwise the evidence is total conflicting, and the rule is not applicable). Furthermore, all of the original bba will be normalized on this common support before being combined. This is the very fundamental principle on which is based DST and the combination of evidence proposed by Shafer.

More precisely in our example, the adjusted bba $m'_2(.)$ of Doctor 2 is described in Tables II- III and IV.

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	b_1
$C \equiv \emptyset$	0	$1-b_1-b_2$
$A \cup B \cup \emptyset = A \cup B$	0	b_2

Table II
STEP 1 OF ADJUSTMENT OF $m_2(.)$.

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	b_1+b_2
$C \equiv \emptyset$	0	$1-b_1-b_2$

Table III
STEP 2 OF ADJUSTMENT OF $m_2(.)$.

Focal elem. \ bba's	$m_1(.)$	$m'_2(.)$
A	a	0
$A \cup B$	$1-a$	$\frac{b_1+b_2}{1-(1-b_1-b_2)} = 1$

Table IV
ADJUSTED AND NORMALIZED BBA'S $m_1(.)$ AND $m'_2(.)$.

After this adjustment, the bba $m'_2(.)$ of Doctor 2 becomes the vacuous bba, which has no impact to the result. This perfectly explains the result produced by DS rule, but doesn't suffice to fully justify its real usefulness for applications.

In general, given two frames of discernment to be combined, if one is a proper subset of the other, the result is asymmetric — the smaller frame always wins the competition, though the other one does not always become vacuous.

Again, here we see that the result is not from any specialty of our emblematic example, but directly from conjunctive nature of the DS rule. As Shafer wrote: "A basic idea of the theory of belief functions is the idea of evidence whose only direct effect on the frame Θ is to support a subset A_1, and an implicit aspect of this idea is that when this evidence is combined with further evidence whose only direct effect on Θ is to establish a compatible subset A_2, the support for A_1 is inherited by $A_1 \cap A_2$." [21]

Now the fundamental question becomes: should evidence combination be treated in this way?

IV. EVALUATING THE VALIDITY OF DST

After sharing the above result we found with other researchers in the field, we got three types of response, which can be roughly categorized as:

1) This result does not show that DST is wrong, but that there are situations where it is not applicable. This example contains conflicting evidence, so DST should not be applied.
2) This result does not show that DST is wrong, and this result is exactly the correct one. It is your intuition that is wrong.
3) This result shows that DST is wrong, since it is unreasonable to let one expert's opinion to completely suppress the other opinions.

The first response is not very satisfactory because it tells us that DST should not be applied when evidences conflict. If we admit such a response, what is the real purpose in using DS rule in practical applications using belief functions, since most of them do involve conflicting sources? In agreeing with the first response, we see that DS rule reduces to the strict conjunctive rule which should be used only in limited cases where there is no conflict between sources. It is not obvious to see why the conjunctive rule even in these cases is well-adapted for the pooling of evidence. In fact, in the context on no conflicting sources, the conjunctive rule corresponds just to the selection of the most specific source, rather than a combination (pooling) of evidences.

Each of the two last responses is supported by a long argument, which sounds reasonable until they are put together — how can we have such different opinions on such a simple example? Can DST be used to combine them to provide a final conclusion based on the pooled evidence?

Instead of trying to apply DS rule (if possible) or to analyze the above responses one by one, we will temporarily step back from this concrete case, and discuss a meta-level problem first, that is, when a mathematical theory is applied to a practical situation, how to decide the validity of this application? In what sense the result is "right" or "wrong"?

Of course, there are some trivial cases where the solution is obvious. If the result is deterministic and there is an objective way to check it, then the conclusion is conclusive. Unfortunately, in the field of uncertain reasoning, it is not that simple. In the above example, we cannot use the disease the patient has (assume we finally become certain about it) to decide whether DST is correctly applied to it, though it may influence our degree of belief about the theory. Actually, this is exactly how "evidence" is different from "proof" in deciding the truthfulness of a conclusion — while a proof can determine the truth-value of a statement *conclusively*, evidence can only do so *tentatively*, because in realistic situations there is always further evidence to come.

Another relatively simple situation is that an internal inconsistency is found in the mathematical theory. In that case

the theory is clearly "wrong", and is not good for any normal usage. This is not the case here, neither. There are inconsistencies founded about DST, such as [11], but it is between the theory and its semantic interpretations (that is, between what it is claimed to do and what it actually does), rather than within the (uninterpreted) mathematical structure of the theory.

What we are facing is a more complicated situation, where the result produced by a theory "sounds wrong", that is, it conflicts with our intuition, experience, or belief. DST is not the only theory that has run into this kind of trouble, and there are indeed three logical possibilities, as represented by the responses listed previously. What to make the situation more complicated is the existence of two types of researchers, with very different motivations in this context:

- **A**: There are people who start with a domain problem, which is called "belief revision", "evidential reasoning", "data fusion", and so on, by different researchers. They are looking for a mathematical tool for this job.
- **B**: There are people who start with a mathematical model that has some properties they like, DST in this case, and are looking for proper practical applications for it.

In general, both motivations are legitimate, but it is crucial that they should not be confused with each other. We belong to Type **A**, and are evaluating DST with respect to the problem we have in mind, to which DST is often claimed to be a solution. For this reason, we argue that DST failed to do the job. Some objection to our conclusion comes from people of Type **B**, to them DST can be called "wrong" only when an internal inconsistency is found, otherwise the theory is always correct, and all mistakes are cased by its human users. Here we are not criticizing DST in that sense. Using the above example, we conclude DST to be "wrong" because it fails to properly handle evidence combination, or in other words, what it claims to do does not match what it actually does, as the defect proved in [11].

To support our conclusion with evidence (rather than with intuition), we start from an analysis of the task of "evidence combination" (or call it "data fusion"). As mentioned above, "evidence" has an impact on "degree of belief" in a system doing evidential reasoning, like "proof" has on "truth-value" in a system using classical logic, except here the impact is tentative and inconclusive (i.e. it doesn't provide an absolute truth). This is exactly why evidence combination becomes necessary (while there is no corresponding operation in classical logic) — in a system that is open to new evidence, it needs to use new evidence to adjust its degree of belief, and the "rule" here should be similar to the rule used to merge the opinions of different experts. In both cases, each opinion has some evidential support, though none of them can be treated as absolutely certain.

This is according to the above understanding of "evidence combination" that DST's result in the above example is considered as "wrong", simple because it allows certain opinion to become immune to revision. To be concrete, what if the previous example consists of 100 doctors, and all of them, except Doctor 1, consider C the most likely disease the patient has, though they cannot completely rule out the possibility of A and B. On the other hand, Doctor 1, for some unspecified reason, considers C impossible, and A more likely than B. In this case, DST will still completely accept Doctor 1's opinion, and ignore the judgment of the other 99 experts. We don't believe anyone will consider this judgment reasonable.

Based on conjunction, DS rule supports the dictatorial power of a source, by accepting the minority opinion as effective solution for "pooling" evidences, no matter that the general a priori assumption applying DS rule is all sources of information are equally reliable, which means all sources' opinions should be taken into account on equal terms. From a theoretical point of view, we don't think this type of belief should be allowed in evidential reasoning; from a practical point of view, such a treatment can lead to serious consequences, since it means that some errors in one evidence channel cannot be corrected by other channels, no matter how many and how strong.

To us, the only possible way to justify DST in similar situations is to change what we mean by "evidence combination". According to Shafer's treatment, "evidence combination" becomes a process similar to constraint satisfaction, where each piece of evidence put some absolute restriction on where the final result can be, and their combination corresponds to "to reach a consensus by mutual constraining". According to this interpretation, Doctor 1 has the right to suppress all the other opinions and therefore can dictates his opinion. If we want to consider each doctor's opinion as absolute truth (following Shafer's interpretation), though sometimes underspecified, then the result becomes acceptable. But in this case, the validity and usefulness of DS rule is strongly conditioned by the justification of the fact that each doctor does really have access to the absolute truth on the proposition under consideration. How can this be done in practice? From what knowledge can a doctor get an absolute truth on a proposition? The answers to these very important questions for validating DS rule haven't been given in the literature so far (to the authors knowledge).

Furthermore, if every doctor is allowed to claim this kind of absolute truth, there is nothing preventing different doctors from announcing different "truths", which leads to "total conflict" situation that cannot be resolved by Dempster's rule. Therefore, the theory faces a paradox: it must either ban the claim of any unrevisable belief, or find a way to handle the conflict among such beliefs. To accept unrevisable beliefs only from a single source does not sound reasonable.

The difference between the two interpretations of "evidence combination" are semantic and philosophical. According to our interpretation, when there are competing opinions supported by distinct evidences, none of them has "absolute truth", but each has some "relative truth", with respect to the supporting evidence, so in the combination process all the opinions can be more or less revised, and the result is usually a compromise; According to Shafer's interpretation, if one source considers an element in the frame of discernment

as impossible, this judgment will be taken as absolute truth, and is therefore unrevisable by the other opinions.

Though it is possible to imagine certain situations, such as Shafer's "random coding" scenario [21], where DST can produce reasonable results, we believe our interpretation of "evidence combination" better matches the common sense meaning of the phrase, as well as the most practical needs in this domain.

It is true that every mathematical theory has its limited applicable domain, and we are not demanding DST to be "universal". However, here the situation is that DST is often presented as a general mechanism for evidential reasoning. Even though it has been widely acknowledged in the community that DST cannot properly handle (highly) conflicting evidence, its cause has not been clearly analyzed, nor is the applicable situations of the theory clearly specified. The above analysis answers these questions: conflicting evidence (whatever they are, in high or in low conflict) cannot be handled well by DST, since they cannot be seen as "partial truth" anymore.

The last important point to underline is the about DS conditioning rule (4) and the formula (5) for conditional plausibility. Let consider Θ and two bba's $m_1(.)$ and $m_2(.)$ defined on 2^Θ and their DS combination $m_{DS}(.) = [m_1 \oplus m_2](.)$ and let assume a conditioning element $Z \neq \emptyset$ in 2^Θ and the bba $m_Z(Z) = 1$, then $m_{DS}(.|Z) = [m_{DS} \oplus m_Z](.) = [m_1 \oplus m_2 \oplus m_Z](.)$. Because $m_{DS}(.) = [m_1 \oplus m_2](.)$ is inconsistent with the probability calculus [10], [11], [14], [15], [20], then $m_{DS}(.|Z)$ is also inconsistent. Therefore for any X in 2^Θ, the conditional plausibility $Pl(X|Z)$ expressed by $Pl(X|Z) = Pl(X \cap Z)/Pl(Z)$ (with apparent similarity with Bayes formula) obtained from $m_{DS}(.|Z)$ is not compatible with the conditional probability as soon as several sources of evidences are involved.

V. Conclusions

In this paper, through a very simple example, we have shown and explained what we consider as a very serious flaw of DS reasoning, which has generated strong controversies in the last three decades. The problem is: given the mathematical property of the combination rule, in certain situation the judgment expressed by a single information source will be effectively treated as absolute truth that will dominate the final result, no matter what judgments the other sources have. Such a result is in total disagreement with the common-sense notion of "evidence combination", "information fusing", or whatever the process is called, because in such a process, each information or evidence source should always be considered only as having local or relative truth. In summary, we believe DST has been often and widely used in situations where it should not, and such applications are wrong. After several decades of existence, proponents of DST need to clearly identify the situations where its model may be truly applicable and what real experimental "pooling" of evidence process DS rule corresponds to. This question is not what this paper is discussing, but is left for future research and discussions.

References

[1] Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton, 1976.

[2] Dempster, A.: Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist., Vol. 38, 325–339, 1967.

[3] Dempster, A.: A generalization of Bayesian inference, J. R. Stat. Soc. B 30, 205–247, 1968.

[4] Dempster, A.: The Dempster-Shafer calculus for statisticians, IJAR, Vol. 48, 365–377, 2008.

[5] Smets, P.: Practical uses of belief functions. in K. B. Lskey and H. Prade, Editors, Uncertainty in Artificial Intelligence 15 (UAI 99), pp. 612–621, Stockholm, Sweden, 1999.

[6] Zadeh, L.A.: On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, U.S.A., 1979.

[7] Zadeh, L.A.: Book review: A mathematical theory of evidence, The Al Magazine, Vol. 5 (3), pp. 81-83, 1984.

[8] Zadeh, L.A.: A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The Al Magazine, Vol. 7 (2), pp. 85–90, 1986.

[9] Lemmer J. : Confidence factors, empiricism and the Dempster-Shafer theory of evidence, in Proc. of 1st Conf. on Uncertainty in Artificial Intelligence (UAI-85), pp. 160–176, 1985.

[10] Voorbraak, F. : On the justification of Demspster's rule of combination, Dept. of Philosophy, Univ. of Utrecht, The Netherlands, Logic Group Preprint Series, No. 42, Dec. 1988.

[11] Wang, P. : A defect in Dempster-Shafer theory, in Proc. of 10th Conf. on Uncertainty in AI, pp. 560–566, 1994.

[12] Pearl, J.: Reasoning with belief functions: An analysis of compatibility, International Journal of Approximate Reasoning, Vol. 4, pp. 363–389, 1990.

[13] Pearl, J.: Rejoinder of comments on "Reasoning with belief functions: An analysis of compatibility", International Journal of Approximate Reasoning, Vol. 6, pp. 425–443, 1992.

[14] Walley, P. : Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, pp. 278–281, 1991.

[15] Gelman, A.: The boxer, the wrestler, and the coin flip: a paradox of robust Bayesian inference and belief functions, American Statistician, Vol. 60, No. 2, pp. 146–150, 2006.

[16] Smarandache, F., Dezert, J.: Advances and applications of DSmT for information fusion, Volumes 1, 2 & 3, ARP, 2004–2009. http://www.gallup.unm.edu/~smarandache/DSmT.htm

[17] Smets, P., Kennes, R. : The transferable belief model, Artif. Int., Vol. 66, pp. 191–234, 1994.

[18] Smets, P. : The transferable belief model for quantified belief representation, Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol.1, Kluwer, 1998.

[19] Dezert J., Tchamova A.: On the behavior of Dempster's rule of combination, Presented at the spring school on Belief Functions Theory and Applications (BFTA), Autrans, France, 4–8 April 2011 (http://hal.archives-ouvertes.fr/hal-00577983/).

[20] Dezert J., Tchamova A., Dambreville, F. : On the mathematical theory of evidence and Dempster's rule of combination, May 2011(http://hal.archives-ouvertes.fr/hal-00591633/fr/).

[21] Shafer, G.: Constructive probability, Synthese, 48(1):1–60, 1981.

Hierarchical Proportional Redistribution for bba Approximation

Jean Dezert
Deqiang Han
Zhunga Liu
Jean-Marc Tacnet

Originally published as Dezert J., Han D., Liu Z., Tacnet J.-M., *Hierarchical proportional redistribution for bba approximation*, in Belief 2012, Compiègne, May 2012, and reprinted with permission.

Abstract. Dempster's rule of combination is commonly used in the field of information fusion when dealing with belief functions. However, it generally requires a high computational cost. To reduce it, a basic belief assignment (bba) approxima-tion is needed. In this paper we present a new bba approximation approach called hierarchical proportional redistribution (HPR) allowing to approximate a bba at any given level of non-specificity. Two examples are given to show how our new HPR works.

1 Introduction

Dempster-Shafer Theory (DST), also called Theory of Evidence [10], has been widely used in many applications, e.g., information fusion, pattern recognition and decision making [11]. Although it is appealing in uncertainty modeling, while appearing more controversial for consistent reasoning, the high computational cost remains problematic which is often raised against its use [11]. To resolve such a problem, three major types of approaches have been proposed.

The first is to propose efficient procedures for performing exact computations [1, 8]. The second is composed of Monte-Carlo techniques [9]. The third is to approximate a belief function to a simpler one. The papers of Voorbraak [13], Dubois and Prade [5] are seminal works of this type. Other representative works include $k-l-x$ [3] and k-additive belief function [2, 6]. Denœux uses hierarchical clustering to implement the inner and outer approximation [3].

In this paper, we propose a new method called hierarchical proportional redistribution (HPR) to approximate any general basic belief assignment (bba) at a given level of non-specificity [4], up to the ultimate level 1 corresponding to a Bayesian bba [10]. The level of non-specificity can be controlled by the users through the adjustment of the maximum cardinality of remaining focal elements. For the approximated bba obtained by HPR, the maximum cardinality of the focal elements is k. Thus HPR can be considered as a generalized k-additive belief approximation. Some examples are given to show how our proposed HPR method works, and to compare it with other approximations.

2 Basics of Dempster-Shafer Theory (DST)

In DST [10], the frame of discernment (FoD) is a set Θ of mutual exhaustive and exclusive elements. $m(.) : 2^{\Theta} \rightarrow [0,1]$ is a basic belief assignment (bba), also called mass function, if it satisfies

$$\sum_{A \subseteq \Theta} m(A) = 1, \; m(\emptyset) = 0. \tag{1}$$

Belief function (Bel) and plausibility function (Pl) are defined as

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad \text{and} \quad Pl(A) = \sum_{A \cap B \neq \emptyset} m(B). \tag{2}$$

Suppose that $m_1, m_2, ..., m_n$ are n bba's, then Dempster's rule of combination is defined by

$$m(A) = \begin{cases} 0, & A = \emptyset \\ \dfrac{\sum\limits_{\cap A_i = A} \prod\limits_{1 \leq i \leq n} m_i(A_i)}{\sum\limits_{\cap A_i \neq \emptyset} \prod\limits_{1 \leq i \leq n} m_i(A_i)}, & A \neq \emptyset \end{cases} \tag{3}$$

This rule is used in DST to combine pieces of evidence expressed by bba's. As referred above, Dempster's combination has high computational cost and three types of approaches have been proposed to reduce it. We prefer belief approximation approaches [2, 3, 6, 12] since they both reduce the computational cost of the combination and allow to deal with smaller-size focal elements, which is more intuitive for human to catch the meaning and interpret fusion results [2].

3 Two bba Approximation Approaches

1) $k-l-x$ **approximation:** This was proposed by Tessem [12]. The simplified bba obtained by $k-l-x$ approach satisfies: a) keep no less than k focal elements; b) keep no more than l focal elements; c) the mass assignment to be deleted is no greater than x. In $k-l-x$, the focal elements of a original bba are sorted by their masses. Such an algorithm chooses the first p focal elements such that $k \leq p \leq l$ and such

that the sum of the masses of these first p focal elements is no less than $1-x$. The deleted masses are redistributed to the other focal elements through a normalization.

2) k-additive belief function approximation: Given $m(.): 2^{\Theta} \rightarrow [0,1]$, one kind of k-additive belief function [2, 6] induced by the mass $m(.)$ is defined by

$$\begin{cases} m_k(B) = m(B) + \sum\limits_{A \supset B, A \subseteq \Theta, |A|>k} \frac{m(A)\cdot|B|}{\mathscr{N}(|A|,k)}, & \forall\, |B| \leq k \\ m_k(B) = 0, & \forall\, |B| > k \end{cases} \tag{4}$$

where $B \subseteq \Theta$ and

$$\mathscr{N}(|A|,k) = \sum_{j=1}^{k} \binom{|A|}{j} \cdot j = \sum_{j=1}^{k} \frac{|A|!}{(j-1)!(|A|-j)!} \tag{5}$$

is the average cardinality of the subsets of A of size at most k. For k-additive belief approximation, the maximum cardinality of available focal elements is no greater than k. Other bba approximation methods can be found in related references.

4 Hierarchical Proportional Redistribution Approximation

In this paper we propose a new bba approximation approach called hierarchical proportional redistribution (HPR), which provides a new way to reduce step-by-step the mass committed to uncertainties. Ultimately an approximate measure of subjective probability can be obtained if needed, i.e. a so-called Bayesian bba in [10]. Our proposed procedure can be stopped at any step in the process and thus it allows to reduce the number of focal elements of a given bba in a simple manner to diminish the size of the core [10] of a bba. Thus we can reduce the complexity (if needed) when applying also some complex rules of combinations. By using HPR, we can obtain approximate bba's at any different non-specificity level that we want. Let us first introduce two new notations for convenience and conciseness:

1. Any element of cardinality $1 \leq k \leq n$ of the power set 2^{Θ} will be denoted $X(k)$ by convention. For example, if $\Theta = \{A,B,C\}$, then $X(2)$ can denote the following partial uncertainties $A \cup B$, $A \cup C$ or $B \cup C$, and $X(3)$ denotes the total uncertainty $A \cup B \cup C$.
2. The proportional redistribution factor (ratio) of width s involving elements X and Y of the powerset is defined by (for $X \neq \emptyset$ and $Y \neq \emptyset$)

$$R_s(Y,X) \triangleq \frac{m(Y) + \varepsilon \cdot |X|}{\sum\limits_{\substack{Y \subset X \\ |X|-|Y|=s}} m(Y) + \varepsilon \cdot |X|} \tag{6}$$

where ε is a small positive number introduced here to deal with particular cases where $\sum\limits_{\substack{Y \subset X \\ |X|-|Y|=s}} m(Y) = 0$.

By convention, we will denote $R(Y,X) \triangleq R_1(Y,X)$ when we use the proportional redistribution factors of width $s=1$, as we use in this paper for this HPR method.

The HPR is a step-by-step (recursive) proportional redistribution of the mass $m(X(k))$ of a given uncertainty $X(k)$ (partial or total) of cardinality $2 \leq k \leq n$ to all the least specific elements of cardinality $k-1$, i.e., to all possible $X(k-1)$, until $k=2$ is reached. The proportional redistribution is done from the masses of belief committed to $X(k-1)$ as done classically in DSmP transformation. The "hierarchical" masses $m_h(.)$ are recursively (backward) computed as follows. Here $m_{h(k)}$ represents the approximate bba obtained at the step $n-k$ of HPR, i.e., it has the maximum focal element cardinality of k.

$$\begin{aligned} &m_{h(n-1)}(X(n-1)) = m(X(n-1)) + \sum_{\substack{X(n)\supset X(n-1),\\ X(n),X(n-1)\in 2^\Theta}} [m(X(n))\cdot R(X(n-1),X(n))]; \\ &m_{h(n-1)}(A) = m(A), \forall |A| < n-1 \end{aligned} \tag{7}$$

$m_{h(n-1)}(\cdot)$ is the bba obtained at the first step of HPR ($n-(n-1)=1$), the maximum focal element cardinality of $m_{h(n-1)}$ is $n-1$.

$$\begin{aligned} m_{h(n-2)}(X(n-2)) &= m(X(n-2)) \\ &\quad + \sum_{\substack{X(n-1)\supset X(n-2)\\ X(n-2),X(n-1)\in 2^\Theta}} [m_{h(n-1)}(X(n-1))\cdot R(X(n-2),X(n-1))] \\ m_{h(n-2)}(A) &= m_{h(n-1)}(A), \forall |A| < n-2 \end{aligned} \tag{8}$$

$m_{h(n-2)}(\cdot)$ is the bba obtained at the second step of HPR ($n-(n-2)=2$), the maximum focal element cardinality of $m_{h(n-2)}$ is $n-2$.

This hierarchical proportional redistribution process can be applied similarly (if one wants) to compute $m_{h(n-3)}(.)$, $m_{h(n-4)}(.)$, ..., $m_{h(2)}(\cdot)$, $m_{h(1)}(\cdot)$ with

$$\begin{aligned} &m_{h(2)}(X(2)) = m(X(2)) + \sum_{\substack{X(3)\supset X(2)\\ X(3),X(2)\in 2^\Theta}} [m_{h(3)}(X(3))\cdot R(X(2),X(3))] \\ &m_{h(2)}(A) = m_{h(3)}(A), \forall |A| < n-2 \end{aligned} \tag{9}$$

$m_{h(2)}(\cdot)$ is the bba obtained at the first step of HPR ($n-2$), the maximum focal element cardinality of $m_{h(2)}$ is 2.

Mathematically, for any $X(1)\in\Theta$, i.e. any $\theta_i\in\Theta$ a Bayesian belief function can be obtained by HPR method in deriving all possible steps of proportional redistributions of partial ignorances in order to get

$$m_{h(1)}(X(1)) = m(X(1)) + \sum_{\substack{X(2)\supset X(1)\\ X(1),X(2)\in 2^\Theta}} [m_{h(2)}(X(2))\cdot R(X(1),X(2))] \tag{10}$$

In fact, $m_{h(1)}(\cdot)$ is a probability transformation, called here the Hierarchical DSmP (HDSmP). Since $X(n)$ is unique and corresponds only to the full ignorance $\theta_1\cup\theta_2\cup\ldots\cup\theta_n$, the expression of $m_h(X(n-1))$ in Eq.(9) just simplifies as

$$m_{h(n-1)}(X(n-1)) = m_h(X(n-1)) + m(X(n))\cdot R(X(n-1),X(n)) \tag{11}$$

For the full proportional redistribution of the masses of uncertainties to the elements least specific involved in these uncertainties, no mass is lost during the step-by-step hierarchical process and thus at any step of HPR, the sum of masses is kept to one.

5 Examples

5.1 Example 1

Let's consider the following bba over $\Theta = \{\theta_1, \theta_2, \theta_3\}$:

$$m(\theta_1) = 0.10, \quad m(\theta_2) = 0.17, \quad m(\theta_3) = 0.03, \quad m(\theta_1 \cup \theta_2) = 0.15,$$
$$m(\theta_1 \cup \theta_3) = 0.20, \quad m(\theta_2 \cup \theta_3) = 0.05, \quad m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30.$$

We apply the HPR with $\varepsilon = 0$ in this example because there is no mass of belief equal to zero. It can be verified that the result obtained with small positive ε parameter remains (as expected) numerically very close to what is obtained with $\varepsilon = 0$.

• **Step 1:** The first step of HPR consists in redistributing back $m(\theta_1 \cup \theta_2 \cup \theta_3) = 0.30$ committed to the full ignorance to the elements $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ only, because these elements are the only elements of cardinality 2 that are included in $\theta_1 \cup \theta_2 \cup \theta_3$. Applying the Eq. (8) with $n = 3$, one gets when $X(2) = \theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_1 \cup \theta_2$ the following masses.

$$m_{h(2)}(\theta_1 \cup \theta_2) = m(\theta_1 \cup \theta_2) + m(X(3)) \cdot R(\theta_1 \cup \theta_2, X(3)) = 0.15 + (0.30 \cdot 0.375) = 0.2625$$

because $R(\theta_1 \cup \theta_2, X(3)) = \frac{0.15}{0.15+0.20+0.05} = 0.375$.
Similarly, one gets

$$m_{h(2)}(\theta_1 \cup \theta_3) = m(\theta_1 \cup \theta_3) + m(X(3)) \cdot R(\theta_1 \cup \theta_3, X(3)) = 0.20 + (0.30 \cdot 0.5) = 0.35$$

because $R(\theta_1 \cup \theta_3, X(3)) = \frac{0.20}{0.15+0.20+0.05} = 0.5$, and also

$$m_{h(2)}(\theta_2 \cup \theta_3) = m(\theta_2 \cup \theta_3) + m(X(3)) \cdot R(\theta_2 \cup \theta_3, X(3)) = 0.05 + (0.30 \cdot 0.125) = 0.0875$$

because $R(\theta_2 \cup \theta_3, X(3)) = \frac{0.05}{0.15+0.20+0.05} = 0.125$.

• **Step 2** Now, we go to the next step of HPR principle and one needs to redistribute the masses of partial ignorances $X(2)$ corresponding to $\theta_1 \cup \theta_2$, $\theta_1 \cup \theta_3$ and $\theta_2 \cup \theta_3$ back to the singleton elements $X(1)$ corresponding to θ_1, θ_2 and θ_3. We use Eq. (10) for doing this as follows:

$$\begin{aligned} m_{h(1)}(\theta_1) &= m(\theta_1) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_1, \theta_1 \cup \theta_2) \quad + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_1, \theta_1 \cup \theta_3) \\ &\approx 0.10 + (0.2625 \cdot 0.3703) + (0.35 \cdot 0.7692) = 0.10 + 0.0972 + 0.2692 = 0.4664 \end{aligned}$$

because $R(\theta_1, \theta_1 \cup \theta_2) = \frac{0.10}{0.10+0.17} \approx 0.3703$ and $R(\theta_1, \theta_1 \cup \theta_3) = \frac{0.10}{0.10+0.03} \approx 0.7692$
Similarly, one gets

$$\begin{aligned} m_{h(1)}(\theta_2) &= m(\theta_2) + m_h(\theta_1 \cup \theta_2) \cdot R(\theta_2, \theta_1 \cup \theta_2) + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_2, \theta_2 \cup \theta_3) \\ &\approx 0.10 + (0.2625 \cdot 0.6297) + (0.0875 \cdot 0.85) = 0.17 + 0.1653 + 0.0744 = 0.4097 \end{aligned}$$

because $R(\theta_2, \theta_1 \cup \theta_2) = \frac{0.17}{0.10+0.17} \approx 0.6297$ and $R(\theta_2, \theta_2 \cup \theta_3) = \frac{0.17}{0.17+0.03} = 0.85$. and also

$$\begin{aligned} m_{h(1)}(\theta_3) &= m(\theta_3) + m_h(\theta_1 \cup \theta_3) \cdot R(\theta_3, \theta_1 \cup \theta_3) + m_h(\theta_2 \cup \theta_3) \cdot R(\theta_3, \theta_2 \cup \theta_3) \\ &\approx 0.03 + (0.35 \cdot 0.2307) + (0.0875 \cdot 0.15) = 0.03 + 0.0808 + 0.0131 = 0.1239 \end{aligned}$$

because $R(\theta_3,\theta_1\cup\theta_3)=\frac{0.03}{0.10+0.03}\approx 0.2307$ and $R(\theta_3,\theta_2\cup\theta_3)=\frac{0.03}{0.17+0.03}=0.15$ Hence, the result of final step of HPR is:

$$m_{h(1)}(\theta_1)=0.4664,\quad m_{h(1)}(\theta_2)=0.4097,\quad m_{h(1)}(\theta_3)=0.1239.$$

We can easily verify that $m_{h(1)}(\theta_1)+m_{h(1)}(\theta_2)+m_{h(1)}(\theta_3)=1$.

To compare HPR with the approach of $k-l-x$, we set the parameters of $k-l-x$ to obtain bba's with equal focal element number with HPR at each step. In Example 1, for HPR at first step, it can obtain a bba with 6 focal elements. Thus we set $k=l=6,x=0.4$ for $k-l-x$ to obtain a bba with 6 focal elements. Similarly, for HPR at second step, it can obtain a bba with 3 focal elements. Thus we set $k=l=3,x=0.4$ for $k-l-x$. Based on HPR and $k-l-x$, the results are shown in Table 1.

Table 1 Experimental results of Example 1.

Focal elements	$m_{h(k)}(\cdot)$ - approximate bba			$m(\cdot)$ obtained by $k-l-x$	
	$k=3$	$k=2$	$k=1$	$k=l=6$	$k=l=3$
θ_1	0.1000	0.1000	0.4664	0.1031	0.0000
θ_2	0.1700	0.1700	0.4097	0.1753	0.2573
θ_3	0.0300	0.0300	0.1239	0.0000	0.0000
$\theta_1\cup\theta_2$	0.1500	0.2625	0.0000	0.1546	0.0000
$\theta_1\cup\theta_3$	0.2000	0.3500	0.0000	0.2062	0.2985
$\theta_2\cup\theta_3$	0.0500	0.0875	0.0000	0.0515	0.0000
$\theta_1\cup\theta_2\cup\theta_3$	0.3000	0.0000	0.0000	0.3093	0.4478

5.2 Example 2

Let's consider $\Theta=\{\theta_1,\theta_2,\theta_3\}$, and the bba $m(\theta_3)=0.7$ and $m(\theta_1\cup\theta_2\cup\theta_3)=0.30$. Here, the masses of all the focal elements with cardinality size 2 equal to zero. For HPR, when $\varepsilon>0$, $m(\theta_1\cup\theta_2\cup\theta_3)$ will be divided equally and redistributed to $\{\theta_1\cup\theta_2\}$, $\{\theta_1\cup\theta_3\}$ and $\{\theta_2\cup\theta_3\}$. Because the ratios are (taking for example $\varepsilon=0.001$)

$$R(\theta_1\cup\theta_2,X(3))=R(\theta_1\cup\theta_3,X(3))=R(\theta_2\cup\theta_3,X(3))=\frac{0.00+0.001\cdot 3}{(0.00+0.001\cdot 3)\cdot 3}=0.3333$$

In this case, HPR cannot work directly when $\varepsilon=0$. This shows the necessity for the use of $\varepsilon>0$. The bba's obtained from $\text{HPR}_{\varepsilon=0.001}$ and $k-l-x$ are listed in Table 2.

From the results of Examples 1 & 2, we can see that based on $k-l-x$, the users can control the number of focal elements but cannot control the maximum cardinality of focal elements. Although based on $k-l-x$, the number of focal elements can be reduced, the focal elements with big cardinality might also be kept. This is not good for further reducing computational cost. But with the proposed HPR method, users can easily control both the non-specificity of approximated bba's and the focal element's size.

Table 2 Experimental results of Example 2 ($\varepsilon = 0.001$)

Focal elements	$m_{h(k)}(\cdot)$ - approximate bba			$m(\cdot)$ obtained by $k-l-x$	
	$k=3$	$k=2$	$k=1$	$k=l=6$	$k=l=3$
θ_1	0.0000	0.0000	0.0503	0.0000	0.0000
θ_2	0.0000	0.0000	0.0503	0.0000	0.0000
θ_3	0.7000	0.7000	0.8994	0.7000	0.7000
$\theta_1 \cup \theta_2$	0.0000	0.1000	0.0000	0.0000	0.0000
$\theta_1 \cup \theta_3$	0.0000	0.1000	0.0000	0.0000	0.0000
$\theta_2 \cup \theta_3$	0.0000	0.1000	0.0000	0.0000	0.0000
$\theta_1 \cup \theta_2 \cup \theta_3$	0.3000	0.0000	0.0000	0.3000	0.3000

5.3 Example 3

In this work, an approximation method 1 (giving $m_1(.)$) is considered better than a method 2 (giving $m_2(.)$) if both conditions are fulfilled: 1) if the distance between $m_1(.)$ and original bba $m(.)$ is smaller than the distance between $m_2(.)$ and original bba $m(.)$, i.e. $d(m_1,m) < d(m_2,m)$; 2) if the approximate non-specificity value $U(m_1)$ is closer (and lower) to the true non-specificity value $U(m)$ than $U(m_2)$. We have used Jousselme's distance [7] which has been proved recently to be a strict distance metric because it is commonly used in applications. The Non-specificity [4] is given by $U(m) = \sum_{A \subseteq \Theta} m(A) \log_2 |A|$. In this example, we make a comparison between HPR (method 1) and k-additive approach (method 2). We have taken $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5\}$ and generated randomly 30 bba's using the algorithm given in [7]. We compute and plot $d(m_1,m)$, $d(m_2,m)$, $U(m)$, $U(m_1)$ and $U(m_2)$ for several levels of approximation. The results are shown in Fig. 1 and indicate clearly the superiority of HPR over the k-additive approach.

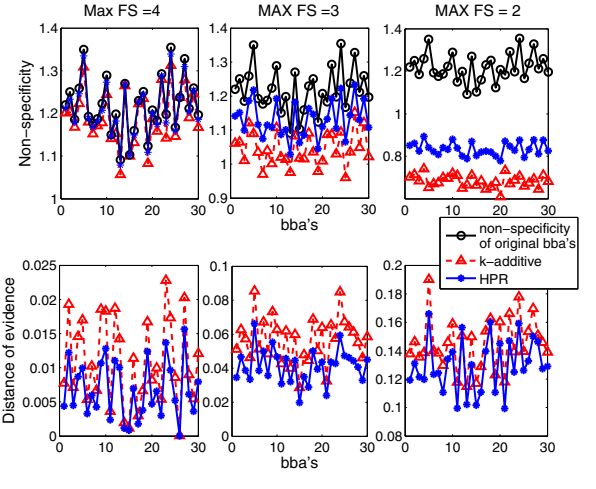


Fig. 1 Results for the Example 3. Comparison of k-additive belief function approximation with HPR approximation method. (FS means Focal element Size)

6 Conclusions

In this paper, a novel bba approximation called HPR has been proposed as an interesting alternative approach to two classical ones. With this HPR, the non-specificity degree can be easily controlled by the users. Our example show its behavior and advantage in comparisons with other well-known bba approximation approaches. HPR has a low computational cost compared with k-additive approach, which will be discussed in a more detailed paper in future. In further works, we will also compare our proposed HPR with more bba approximation approaches available in the literature. In this paper, we have used only the distance of evidence and the non-specificity criteria, which in fact are not enough, or comprehensive to evaluate efficiently bba approximations. So in future, we will try to propose more efficient evaluation criteria to evaluate and design better bba approximations (if possible).

References

1. Barnett, J.A.: Computational methods for a mathematical theory of evidence. In: Proceedings of IJCAI 1981, Vancouver, pp. 868–875 (1981)
2. Burger, T., Cuzzolin, F.: The barycenters of the k-additive dominating belief functions and the pignistic k-additive belief functions. In: Workshop on Theory of Belief Functions, Brest, France, pp. 1–6 (2010)
3. Denœux, T.: Inner and outer approximation of belief structures using a hierarchical clustering approach. International Journal of Uncertainty, Fuzziness, and Knowledge-based Systems 9, 437–460 (2001)
4. Dubois, D., Prade, H.: A note on measures of specificity for fuzzy sets. International Journal of General Systems 10, 279–283 (1985)
5. Dubois, D., Prade, H.: An alternative approach to the handling of subnormal possiblity distributions. Fuzzy Sets and Systems 24, 123–126 (1987)
6. Grabisch, M.: Upper approximation of non-additive measures by k-additive measures - the case of belief functions. In: Proc. of the 1st Int. Symposium on Imprecise Probabilities and Their Applications, Ghent, Belgium (1999)
7. Jousselme, A.-L., Maupin, P.: Distances in evidence theory: Comprehensive survey and generalizations. International Journal of Approximate Reasoning 53, 118–145 (2011)
8. Kennes, R.: Computational aspects of the M*ö*bius transform of graphs. IEEE Transactions on SMC 22, 201–223 (1992)
9. Moral, S., Salmerón, A.: A Monte Carlo Algorithm for Combining Dempster-Shafer Belief Based on Approximate Pre-computation. In: Hunter, A., Parsons, S. (eds.) ECSQARU 1999. LNCS (LNAI), vol. 1638, pp. 305–315. Springer, Heidelberg (1999)
10. Shafer, G.: A Mathematical Theory of Evidence. Princeton University, Princeton (1976)
11. Smets, P.: Practical uses of belief functions. In: Lskey, K.B., Prade, H. (eds.) Uncertainty in Artificial Intelligence 15 (UAI 1999), Stockholm, Sweden, pp. 612–621 (1999)
12. Tessem, B.: Approximations for efficient computation in the theory of evidence. Artificial Intelligence 61, 315–329 (1993)
13. Voorbraak, F.: A computationally efficient approximation of Dempster-Shafer theory. Int. J. Man-Machine Studies 30, 525–536 (1989)

On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory

Albena Tchamova
Jean Dezert

Originally published as Tchamova A., Dezert J., *On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory,* IEEE IS'2012, Sofia, Bulgaria, Sept. 6-8, 2012 (Best paper awards), and reprinted with permission.

***Abstract*—On the base of simple emblematic example we analyze and explain the inconsistent and inadequate behavior of Dempster-Shafer's rule of combination as a valid method to combine sources of evidences. We identify the cause and the effect of the dictatorial power behavior of this rule and of its impossibility to manage the conflicts between the sources. For a comparison purpose, we present the respective solution obtained by the more efficient PCR5 fusion rule proposed originally in Dezert-Smarandache Theory framework. Finally, we identify and prove the inherent contradiction of Dempster-Shafer Theory foundations.**

***Keywords*—Belief functions; Dempster-Shafer Theory; DSmT; PCR5; contradiction.**

I. Introduction

Dempster-Shafer Theory (DST), also known as the Theory of Evidence or the Theory of Belief Functions, was introduced by Shafer in 1976 [1] based on Dempster's previous works [2], [3], [4]. This theory offers an elegant theoretical framework for modeling uncertainty, and provides a method for combining distinct bodies of evidence collected from different sources. In the past more than three decades, DST has been used in many applications, in fields including information fusion, pattern recognition, decision making [5], etc.

In spite of it, starting from Zadeh's criticism [6], [7], [8], many questions have arisen about the validity and the consistency of this theory when combining uncertain and conflicting evidences expressed as basic belief assignments (bba's). Besides Zadeh's example, there have been several detailed analyses on this topic by Lemmer [9], Voorbraak [10] and Wang [11]. Other authors like Pearl [12] and Walley [13], and more recently Gelman [14], have also warned the "belief function community" about this fundamental problem, i.e., the validity of Dempster-Shafer's rule[1] (DS rule for short) for combining distinct pieces of evidences. Since the mid-1990's, many researchers and engineers working with belief functions in applications have observed and admitted that DS rule is problematic for evidence combination, specially when the sources of evidence are highly conflicting.

In response to this challenge, various attempts have been made to circumvent the counter-intuitive behaviors of DS rule. They either replace Dempster-Shafer's rule by alternative rules, listed for example in [15] (Vol. 1), or apply novel semantic interpretations to the functions [15], [16], [17]. This work is based on preliminary ideas presented in the Spring School on Belief Functions Theory and Applications in April 2011 [18], and on many fruitful discussions with colleagues using belief functions. We start from a very basic, but emblematic example to show what is really questionable in DS rule. We demonstrate that the main problem applying DS rule comes not from the level of conflict between sources to combine, but from the underlying interpretation of evidence and degree of belief on which the combination rule is based. We make a comparison with respective results, obtained by using Proportional Conflict Redistribution rule no.5 (PCR5) defined within Dezert-Smarandache Theory (DSmT) [15]. In Section II we briefly recall basics of DST and DS rule. Basics of PCR5 fusion rule are outlined in Section III. In Section IV we describe our basic example and discuss the counter-intuitive result obtained by DS rule and its strange behavior corresponding to the dictatorial power of particular source of evidence with respect to all another sources. A comparison with respective results obtained by PCR5 fusion rule is also made. After a discussion on dictatorial power of DS rule in Section V, we establish and prove in Section VI a fundamental theorem on the contradiction, grounded in DST foundations. Concluding remarks are given in Section VII.

II. Basics of DST

Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ be a frame of discernment of a problem under consideration containing n distinct elements θ_i, $i = 1, \ldots, n$. A basic belief assignment (bba, also called a belief mass function) $m(.) : 2^\Theta \to [0, 1]$ is a mapping from the power set of Θ (i.e. the set of subsets of Θ), denoted 2^Θ, to $[0, 1]$, that must satisfy the following conditions: 1) $m(\emptyset) = 0$, i.e. the mass of empty set (impossible event) is zero; 2)

[1]This rule is also called Dempster's rule in the literature because it was originally proposed by Dempster. We prefer to name it *Dempster-Shafer's rule* because it has widely been promoted by Shafer in his development of theory of belief functions (a.k.a. DST).

$\sum_{X\in 2^\Theta} m(X) = 1$, i.e. the mass of belief is normalized to one. The quantity $m(X)$ represents the mass of belief exactly committed to X. An element $X \in 2^\Theta$ is called a focal element if and only if $m(X) > 0$. The set $\mathcal{F}(m) \triangleq \{X \in 2^\Theta | m(X) > 0\}$ of all focal elements of a bba $m(.)$ is called the core of the bba. By definition, a Bayesian bba $m(.)$ is a bba having only focal elements of cardinality 1. The vacuous bba characterizing full ignorance is defined by $m_v(.) : 2^\Theta \to [0;1]$ such that $m_v(X) = 0$ if $X \neq \Theta$, and $m_v(\Theta) = 1$.

From any bba $m(.)$, the belief function $Bel(.)$ and the plausibility function $Pl(.)$ are defined for $\forall X \in 2^\Theta$ as: $Bel(X) = \sum_{Y|Y\subseteq X} m(Y)$ and $Pl(X) = \sum_{Y|X\cap Y\neq\emptyset} m(Y)$. $Bel(X)$ represents the whole mass of belief that comes from all subsets of Θ included in X. It is interpreted as the lower bound of the probability of X, i.e. $P_{\min}(X)$. $Bel(.)$ is a subadditive measure since $\sum_{\theta_i\in\Theta} Bel(\theta_i) \leq 1$. $Pl(X)$ represents the whole mass of belief that comes from all subsets of Θ compatible with X (i.e., those intersecting X). $Pl(X)$ is interpreted as the upper bound of the probability of X, i.e. $P_{\max}(X)$. $Pl(.)$ is a superadditive measure since $\sum_{\theta_i\in\Theta} Pl(\theta_i) \geq 1$. $Bel(X)$ and $Pl(X)$ are classically seen as lower and upper bounds of an unknown probability $P(.)$ and one has the following inequality satisfied $Bel(X) \leq P(X) \leq Pl(X)$, $\forall X \in 2^\Theta$.

The DS rule of combination [1] is a mathematical operation, denoted $\oplus$, which corresponds to the normalized conjunctive fusion rule. Based on Shafer's model of the frame, the combination of two independent and distinct sources of evidences characterized by their bba $m_1(.)$ and $m_2(.)$ and related to the same frame of discernment Θ is defined by $m_{DS}(\emptyset) = 0$, and $\forall X \in 2^\Theta \setminus \{\emptyset\}$ by

$$m_{DS}(X) = [m_1 \oplus m_2](X) = \frac{m_{12}(X)}{1 - K_{12}} \tag{1}$$

where

$$m_{12}(X) \triangleq \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2 = X}} m_1(X_1)m_2(X_2) \tag{2}$$

corresponds to the conjunctive consensus on X between the two sources of evidence. K_{12} is the total *degree of conflict* between the two sources of evidence defined by

$$K_{12} \triangleq m_{12}(\emptyset) = \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2 = \emptyset}} m_1(X_1)m_2(X_2) \tag{3}$$

When $K_{12} = m_{12}(\emptyset) = 1$, the two sources are said to be in total conflict and their combination cannot be applied since DS rule (1) is mathematically not defined because of 0/0 indeterminacy [1]. DS rule is commutative and associative which makes it very attractive from engineering implementation standpoint, since the combinations of sources can be done sequentially instead globally and the order doesn't matter. Moreover, the vacuous bba is a neutral element for the DS rule, i.e. $[m \oplus m_v](.) = [m_v \oplus m](.) = m(.)$ for any bba $m(.)$ defined on 2^Θ which seems to be an expected[2] property, i.e. a full ignorant source doesn't impact the fusion result.

The conditioning of a given bba $m(.)$ by a conditional element $Z \in 2^\Theta \setminus \{\emptyset\}$ has been also proposed by Shafer [1]. This function $m(.|Z)$ is obtained by DS combination of $m(.)$ with the bba $m_Z(.)$ only focused on Z, i.e. such that $m_Z(Z) = 1$. For any element X of the power set 2^Θ this is mathematically expressed by

$$m(X|Z) = [m \oplus m_Z](X) = [m_Z \oplus m](X) \tag{4}$$

It has been proved [1] that this rule of conditioning expressed in terms of plausibility functions yields to the formula

$$Pl(X|Z) = Pl(X \cap Z)/Pl(Z) \tag{5}$$

which is very similar to the well-known Bayes formula $P(X|Z) = P(X \cap Z)/P(Z)$. Because of this, DST has been widely considered as a generalization of Bayesian inference [3], or equivalently, that probability theory is a special case of the Mathematical Theory of Evidence when manipulating Bayesian bba's.

Despite of the appealing properties of DS rule, its apparent similarity with Bayes formula for conditioning, and many attempts to justify its foundations, several challenges on the theory's validity have been put forth in the last decades, and remain unanswered. For instance, an experimental protocol to test DST was proposed by Lemmer in 1985 [9], and his analysis shows an inherent paradox (contradiction) of DST. Following a different approach, an inconsistency in the fundamental postulates of DST was proved by Wang in 1994 [11]. Some other related works have been listed in the introduction of this paper. In Section IV, we show through a basic emblematic example where does the problem of DS rule comes from, and why it is very risky to use it in very sensible applications specially where security, defense and safety are involved. Before this, we just recall in the next section the principle of the Proportional Conflict Redistribution rule no. 5 (PCR5) defined within DSmT framework [15] to combine bba's.

III. Basics of PCR5 fusion rule

The idea behind the Proportional Conflict Redistribution rule no. 5 defined within DSmT [15] (Vol. 2) is to transfer conflicting masses (total or partial) proportionally to non-empty sets involved in the model according to all integrity constraints. The general principle of PCR rules is to: *1)* calculate the conjunctive consensus between the sources of evidences; *2)* calculate the total or partial conflicting masses; *3)* redistribute the conflicting mass (total or partial) proportionally on non-empty sets involved in the model according to all integrity constraints. Under Shafer's model assumption[3] of the frame Θ, the PCR5 combination rule for only two

[2] A discussion on this topic can be found in [19].

[3] We consider only Shafer's model in this paper and in our examples to make the comparison with Dempster-Shafer's rule results.

sources of information is defined as: $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{Y \in 2^\Theta \setminus \{X\} \\ X \cap Y = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)}\right] \quad (6)$$

All sets involved in the formula (6) are in canonical form. $m_{12}(X)$ corresponds to the conjunctive consensus, i.e:

$$m_{12}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2).$$

All denominators are different from zero. If a denominator is zero, that fraction is discarded. No matter how big or small is the conflicting mass, PCR5 mathematically does a better redistribution of the conflicting mass than Dempster-Shafer's rule since PCR5 goes backwards on the tracks of the conjunctive rule and redistributes the partial conflicting masses only to the sets involved in the conflict and proportionally to their masses put in the conflict, considering the conjunctive normal form of the partial conflict. PCR5 is quasi-associative and also preserves the neutral impact of the vacuous belief assignment, but contrariwise to DS rule the PCR5 fusion rule doesn't allow the dictatorial power of a source as it will be shown in Section IV. With PCR5 rule, the fusion result can always be revised as soon as informative evidences (i.e. not vacuous bba's) become available.

IV. An emblematic example showing the dictatorial power of Dempster-Shafer's rule

Here we present an emblematic example showing the inadequate behavior of Dempster-Shafer's rule. We call this behavior the *dictatorial power* (DP) of DS rule realized by a given source, which is fundamental in DS reasoning. This parametric example is not related to the level of conflict between sources. In this example the level of conflict can be chosen at any low or high value. We show clearly that Dempster-Shafer's rule is not responding to the combination of different bba's since it provides always one and the same results which is not a good expected behavior for a good fusion rule for applications corresponding to the classical[4] sense of pooling of evidences [20].

Let's consider the following frame[5] $\Theta = \{A, B, C\}$ with Shafer's model. We consider two bba's listed in the Table I, associated with two distinct bodies of evidence[6] with parameters a, b_1, and b_2 that can take any values, as long as $a \in [0, 1]$, $b_1, b_2 > 0$, and $b_1 + b_2 \in [0, 1]$.

We grant that all the a priori assumptions below, considered in DST are fulfilled:

1) The sources of evidences are independent;

[4]when putting all evidences together.

[5]Θ could correspond by example to three distinct pathologies of a patient.

[6]In a medical context, the two sources of evidences could correspond to two distinct Doctors providing their own medical diagnostics for a same patient.

TABLE I
Input bba's $m_1(.)$ and $m_2(.)$.

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1 - a$	b_1
C	0	$1 - b_1 - b_2$
$A \cup B \cup C$	0	b_2

2) Both of sources are equally reliable, i.e both of them are equally truthful. As an additional third assumption in this parametric example we consider:

3) Both of sources are truly informative hence no one represents a full ignorant source. It means both sources have their own specific opinions about the particular problem under consideration, which should be taken into account into the fusion process in equal rights manner.

When applying DS rule of combination, one gets:

1) using the conjunctive operator:

$$m_{12}(A) = a(b_1 + b_2) \quad (7)$$

$$m_{12}(A \cup B) = (1 - a)(b_1 + b_2) \quad (8)$$

$$K_{12} = m_{12}(\emptyset) = 1 - b_1 - b_2 \quad \text{(conflicting mass)} \quad (9)$$

2) after normalizing by $1 - K_{12} = b_1 + b_2$, the result is :

$$m_{DS}(A) = \frac{m_{12}(A)}{1 - K_{12}} = \frac{a(b_1 + b_2)}{b_1 + b_2} = a = m_1(A) \quad (10)$$

$$m_{DS}(A \cup B) = \frac{m_{12}(A \cup B)}{1 - K_{12}} = \frac{(1 - a)(b_1 + b_2)}{b_1 + b_2} = 1 - a = m_1(A \cup B) \quad (11)$$

The final result obtained by using DS rule shows clearly that:

- Nevertheless the assumption no. 3 is fulfilled for source $m_2(.)$ (it is obviously a truly informative source of evidence), its opinion doesn't count at all in the fusion process, performed by DS rule since one finally gets $m_{DS}(.) = m_1(.)$. It plays in fact a role of full ignorant source, represented by the vacuous belief assignment $m_v(A \cup B \cup C) = 1$, since $m_{DS}(.) = m_1(.)$ in the DST fusion process. It is against the required a priori assumption no. 2 of DST, for equally reliable/truthful sources of evidence with opinions that have to be taken into account in equal terms.
- The level of conflict $K_{12} = 1 - b_1 - b_2$ encountered between the two sources doesn't matter at all in DS fusion process here, since it can be chosen at any level, depending on the choice of $b_1 + b_2$. No matter how high or how low the conflict is, the result remains one and same: $m_{DS}(.) = m_1(.)$.

In clear, the source 1 dictates his opinion through Dempster-Shafer's rule which is what we consider a very inadequate behavior for solving the problem of combination of evidences in practice. Before analyzing this fundamental problem of DST, let's first take the position of *devil's advocate*, and try to defend the legitimacy of DST's behavior. If we fully trust source 1, the hypothesis C must be ruled out of the frame, because $Bel_1(C) = Pl_1(C) = 0$. So, according to source 1, the original frame of discernment $\Theta = \{A, B, C\}$ should be

TABLE II
ADJUSTED INPUT BBA'S (STEP 1).

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	b_1
$C \equiv \emptyset$	0	$1-b_1-b_2$
$A \cup B \cup \emptyset$	0	b_2

TABLE III
ADJUSTED INPUT BBA'S (STEP 2).

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	b_1+b_2
$C \equiv \emptyset$	0	$1-b_1-b_2$

reduced to $\Theta' = \{A, B\}$, because $C \equiv \emptyset$ (based on the report of source 1). If we consider C impossible to occur, then the report (bba) of source 2 must be adapted/revised according to Tables II and III. Because $m_2(.)$ must be a normalized bba, the masses of all focal elements of $m_2(.)$ are divided by $1-m_2(\emptyset) = b_1+b_2$ so that after adjustment and normalization of $m_2(.)$, the two bba's to combine are presented in Table IV. Based on this reasoning, we see that the adjusted and normalized bba $m'_2(.)$ plays indeed the role of the vacuous bba $m_v(.)$ when working with the reduced frame $\Theta' = \{A, B\}$, which perfectly explains the result produced by DS rule.
Such kind of reasoning unfortunately doesn't prove that the result makes sense, nor it is correct. In fact such reasoning shows clearly an asymmetry in the processing, since the source 1 is assumed to provide an absolute certainty on the event "C cannot occur for sure", whereas the source 2 is adjusted (conditioned) by the declaration of source 1. Such *devil's advocate* reasoning is in fact fallacious, totally mistaken and wrong because it erroneously interprets the impossibility of occurrence of C as a definitive absolute truth (as if all knowledge/evidences were available at the source 1) to withdraw the hypothesis C of the original frame Θ. In fact, the impossibility of C must be interpreted only as conditional truth because it is based only on the partial knowledge related to source 1 (and not on the whole knowledge expressed when pooling the evidences of the two sources).

Let's, just for a comparison purpose, present the respective solution of our example, obtained by DSmT based PCR5 fusion rule. The proportional redistribution of the mass of the partial conflict $m_1(A)m_2(C) = a(1-b_1-b_2)$ is done by

$$\frac{x_A}{m_1(A)} = \frac{x_C}{m_2(C)} = \frac{m_1(A)m_2(C)}{m_1(A)+m_2(C)} = \frac{a(1-b_1-b_2)}{a+1-b_1-b_2}$$

hence $x_A = \frac{a^2(1-b_1-b_2)}{a+1-b_1-b_2}$ and $x_C = \frac{a(1-b_1-b_2)^2}{a+1-b_1-b_2}$.

TABLE IV
ADJUSTED BBA'S $m_1(.)$ AND $m'_2(.)$.

Focal elem. \ bba's	$m_1(.)$	$m'_2(.)$
A	a	0
$A \cup B$	$1-a$	$\frac{b_1+b_2}{1-(1-b_1-b_2)} = 1$

Similarly the redistribution of the partial conflict mass $m_1(A \cup B)m_2(C) = (1-a)(1-b_1-b_2)$ is done by

$$\frac{y_{A\cup B}}{m_1(A\cup B)} = \frac{y_C}{m_2(C)} = \frac{m_1(A\cup B)m_2(C)}{m_1(A\cup B)+m_2(C)}$$

hence $y_{A\cup B} = \frac{(1-a)^2(1-b_1-b_2)}{1-a+1-b_1-b_2}$ and $y_C = \frac{(1-a)(1-b_1-b_2)^2}{1-a+1-b_1-b_2}$.

Therefore with PCR5, one gets a fusion result that does react efficiently to the values of all the masses of focal elements of each source since one has:

$$\begin{aligned} m_{PCR5}(A) &= m_{12}(A) + x_A \\ &= a(b_1+b_2) + \frac{a^2(1-b_1-b_2)}{a+1-b_1-b_2} \end{aligned} \quad (12)$$

$$\begin{aligned} m_{PCR5}(A\cup B) &= m_{12}(A\cup B) + y_{A\cup B} \\ &= (1-a)(b_1+b_2) + \frac{(1-a)^2(1-b_1-b_2)}{2-a-b_1-b_2} \end{aligned} \quad (13)$$

$$\begin{aligned} m_{PCR5}(C) &= x_C + y_C \\ &= \frac{a(1-b_1-b_2)^2}{a+1-b_1-b_2} + \frac{(1-a)(1-b_1-b_2)^2}{2-a-b_1-b_2} \end{aligned} \quad (14)$$

In comparison to DS rule performance, the result obtained by using PCR5 rule, shows clearly that PCR5 fusion rule works efficiently in any level of conflict, taking into account all the a priori assumptions $(1-3)$.

V. DISCUSSION AND ANALYSIS

The result obtained by DS rule according to the example in Section IV seriously calls in question DS rule's validity, as well as its applicability in real fusion problems. We claim that such a result is not acceptable at all. This example is more crucial than the examples discussed in the existing literature, because it shows clearly a serious flaw in DST behavior, since in this example the level of conflict between sources doesn't play a role, so that it cannot be argued that in such case DS must not be applied because of high conflicting situation. We can choose a low conflict level and the result is still the same. The problem remains and the DST based result could become a source of dramatical consequences, especially in cases, related to human health or security. We claim that the problem behind DS rule behavior comes not from the level of conflict between the sources, but from something else.

A. The dictatorial power of source's minority opinion

Let's recall again the example, its strange results, and discuss about the reasoning process behind DS rule. The a priori defined finite frame of discernments $\Theta = \{A, B, C\}$ satisfies Shafer's requirement for a set of truly exhaustive and exclusive hypotheses. Lets's first pay attention on the bba associated with source 1. What is obvious and special from $m_1(.)$, it is the fact, that $Pl_1(C) = 0$. One can reason from here as follows:

1) Source 1 rules out with absolute certainty the hypothesis C considering it as impossible, because of $Pl_1(C) = 0$.

TABLE V
INPUT BBA'S $m_1(.)$ AND $m_2(.)$ FOR THE CASE OF TOTAL CONFLICT.

Focal elem. \ bba's	$m_1(.)$	$m_2(.)$
A	a	0
$A \cup B$	$1-a$	0
C	0	1

2) The above opinion of source 1 (hypothesis C considered as absolutely impossible) cannot be revised if new informative evidence is available for fusion. According to Shafer's definition [1], $Pl_1(C) = 0$ means for every $X \in 2^\Theta$ that $X \cap C \neq \emptyset$, $m_1(X) = 0$. When DS rule is applied to combine $m_1(.)$ and an arbitrary $m'(.)$ (in our example $m'(.) = m_2(.)$), for every $Y \in 2^\Theta$ that $Y \cap C \neq \emptyset$, $m_{DS}(Y) = 0$, because it is the sum of some products, each of them take one of the above $m_1(X)$ as a factor. Consequently, $Pl_{DS}(C) = 0$, no matter what the other source of evidence is.

3) Since with DS rule, the source 1 imposes its own opinion on source 2, and in fact on any other sources (as soon as they have a core including the core of source 1), DST supports the dictatorial power of a given source by accepting the minority opinion as a valid solution of the "fusion of evidences", and by banning in the same time all other sources' different opinions. This behavior is in full contradiction with the a priori assumption no. 2 of DST for equally reliable sources of information, which means their opinions should be taken into account on equal terms in the fusion process (see [20] for a complementary analysis).

B. On the total conflict case banned by DST

Let's try to reveal now what is the logic behind the case, that DS rule cannot solve because of the indefiniteness $(0/0)$ - the case of total conflicting sources of information. We consider the same frame of discernments $\Theta = \{A, B, C\}$ and two bba's (listed in a Table V), associated with two distinct bodies of evidence $m_1(.)$ and $m_2(.)$ with parameter $a \in [0, 1]$. It is obvious from Table V that:

1) Source 1 rules out with an absolute certainty hypothesis C considering it as impossible since $Pl_1(C) = 0$.

2) Source 2 rules out with an absolute certainty the hypotheses A and B considering them as impossible since $Pl_2(A \cup B) = 0$.

The a priori DST assumptions $(1-3)$ still hold. So, the question is: *Which source will possess the dictatorial power in this special case?* Following Shafer's interpretation in this example, the answer is: *both of sources have access to the Absolute truth*. But what is paradoxical and contradictory is that having simultaneously an access to the *Absolute truth*, both of sources ban mutually each other opinions.

Therefore Shafer's interpretation that allows both sources to rule out all possible *Absolute truths* in absolute manner leads to the strong contradiction by accepting the assertion that DS rule cannot be used in such totally conflicting case.

This assertion is substantiated on the obtained mathematical indefiniteness $(0/0)$ as impossible "fusion result". But actually behind the formal mathematical explanation, there resides a real and strong logic that Shafer's distinct *Absolute truth* interpretation granted to each source doesn't hold. The *Absolute truth* is unique and it cannot yield to contradictions in the fusion process.

For a comparison purpose, let's again to present the respective solution in this special conflicting case, obtained by DSmT based PCR5 fusion rule. The proportional redistribution of the mass of the partial conflict $m_1(A)m_2(C) = a$ is done by

$$\frac{x_A}{m_1(A)} = \frac{x_C}{m_2(C)} = \frac{m_1(A)m_2(C)}{m_1(A) + m_2(C)} = \frac{a}{1+a}$$

with $x_A = \frac{a^2}{1+a}$ and $x_C = \frac{a}{1+a}$.

Similarly the proportional redistribution of the partial conflict mass $m_1(A \cup B)m_2(C) = (1-a)$ is done by

$$\frac{y_{A\cup B}}{m_1(A \cup B)} = \frac{y_C}{m_2(C)} = \frac{m_1(A \cup B)m_2(C)}{m_1(A \cup B) + m_2(C)}$$

with $y_{A\cup B} = \frac{(1-a)^2}{2-a}$ and $y_C = \frac{(1-a)}{2-a}$.

Finally, one gets using PCR5 fusion rule

$$m_{PCR5}(A) = x_A = \frac{a^2}{1+a} \tag{15}$$

$$m_{PCR5}(A \cup B) = y_{A\cup B} = \frac{(1-a)^2}{2-a} \tag{16}$$

$$m_{PCR5}(C) = x_C + y_C = \frac{a}{1+a} + \frac{(1-a)}{2-a} \tag{17}$$

It is obvious, DSmT based PCR5 fusion rule works efficiently even in this special total conflicting case. This very attractive rule is just a non-Bayesian reasoning approach, which is not based on such inherent contradiction, as DST, because PCR5 doesn't support Shafer's interpretation of source committed *Absolute truth* and doesn't allow dictatorial power of single source opinion on all other sources, involved in the fusion.

C. Remark on Dempster-Shafer conditioning

Some comments must be given also about DS conditioning rule (4) and the expression (5) for the conditional plausibility. Let consider Θ and two bba's $m_1(.)$ and $m_2(.)$ defined on 2^Θ and their DS combination $m_{DS}(.) = [m_1 \oplus m_2](.)$ and let assume a conditioning element $Z \neq \emptyset$ in 2^Θ and the bba $m_Z(Z) = 1$, then

$$m_{DS}(.|Z) = [m_{DS} \oplus m_Z](.) = [m_1 \oplus m_2 \oplus m_Z](.)$$

Because $m_{DS}(.) = [m_1 \oplus m_2](.)$ is inconsistent with the probability calculus [10], [11], [13], [14], [19], then $m_{DS}(.|Z)$ is also inconsistent. Therefore for any X in 2^Θ, the conditional plausibility $Pl(X|Z)$ expressed by $Pl(X|Z) = Pl(X \cap Z)/Pl(Z)$ obtained from $m_{DS}(.|Z)$, having an apparent similarity with Bayes formula, is in fact not compatible with the conditional probability as soon as several sources of evidences are involved.

VI. Fundamental theorem on the inherent contradiction in DST foundations

On the base of the previous examples and after a detailed analysis of results drawn from Dempster-Shafer's rule and DST reasoning discussed in previous section, we establish the fundamental theorem on the inherent contradiction of DST foundations.

Theorem : Dempster-Shafer Theory is wrong because its foundation is based on an inherent logical contradiction.

Proof : In the basis of DST [1], Shafer considers:

- *An a priori defined finite frame of discernment* $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ *with* $n \geq 2$, *satisfying Shafer's requirement for a set of truly exhaustive and exclusive hypotheses. Recalling Shafer's statement about DST [1] (p. 36): "This formalism is most easily introduced in the case where we are concerned with the true value of some quantity. If we denote the quantity by* θ *and the set of its possible values by* Θ, *then the propositions of interest are precisely those of the form "The true value of* θ *is in* T," *where* T *is a subset of* Θ".
- *Available independent sources of evidences associated with corresponding bba's* $m_i(.)$, $i = 1, 2..$, *where all the sources are equally reliable/trustable and can be truly informative (not fully ignorant).*
- *The level of conflict between the sources can take any low or high value strictly less than one to make Dempster-Shafer's rule mathematically defined.*

On the base of above considerations, one encounters the fundamental contradiction:

1) A given source of evidence $m_p(.)$ can become unrevisable during the fusion when it is allowed to rule out with absolute certainty some hypothesis $\theta_k, k \in [1, n]$ in the frame Θ (if $Pl_p(\theta_k) = 0$ as shown in our emblematic example).
2) DS rule cannot solve the case of total conflict between the sources (because of mathematical indefiniteness $0/0$). This corresponds to the case when both sources: *1)* have an access to the *Absolute truth*; *2)* can become unrevisable during the fusion if they allowed to rule out with absolute certainty all hypotheses in the frame Θ, banning mutually each other opinions. The inability of DS rule to solve this case strongly supports the assertion that the *Absolute truth* must be unique. Otherwise the total conflict case could also be solved/processed by DS rule. So, Shafer's interpretation of distinct *Absolute truth* granted to each source does not hold.

Therefore from the point 2), DST agrees with the assertion that the *Absolute truth* is unique and cannot be a contradiction. This assertion is fully contradicting with Shafer's interpretation of distinct *Absolute truth* granted to each source stated in point 1). This proves the fundamental contradiction in the foundations of DST and completes the proof of our Theorem.

VII. Conclusion

In this paper, we have identified and put in light the very serious inherent contradiction of Dempster-Shafer Theory foundations. On the base of simple emblematic example, we have analyzed and explained the inconsistent and inadequate behavior of Dempster-Shafer's rule of combination as a valid method for the combination of sources of evidences. We have identified the cause and the effect of the dictatorial power behavior of this rule and of its impossibility to manage the conflicts between the sources in a consistent logical way. For a comparison purpose, the respective solutions obtained by the more adequate PCR5 fusion rule, proposed originally in Dezert-Smarandache Theory framework, were presented. This very attractive rule is corresponds to a non-Bayesian reasoning approach, which is not based on such inherent contradiction, as DST, because PCR5 doesn't support Shafer's interpretation of source committed *Absolute truth* and doesn't allow dictatorial power of single source opinion on all other sources, involved in the fusion.

References

[1] Shafer, G. A Mathematical Theory of Evidence. Princeton University Press, Princeton, 1976.
[2] Dempster, A. Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist., Vol. 38, 1967, pp. 325–339.
[3] Dempster, A. A generalization of Bayesian inference, J. R. Stat. Soc. B 30, 1968, pp. 205–247.
[4] Dempster, A. The Dempster-Shafer calculus for statisticians, IJAR, Vol. 48, 2008, pp. 365–377.
[5] Smets, P. Practical uses of belief functions. in K. B. Lskey and H. Prade, Editors, UAI 99, Stockholm, Sweden, 1999, pp. 612–621.
[6] Zadeh, L.A. On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, U.S.A., 1979.
[7] Zadeh, L.A. Book review: A mathematical theory of evidence, The Al Magazine, Vol. 5 (3), 1984, pp. 81–83.
[8] Zadeh, L.A. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The Al Magazine, Vol. 7 (2), 1986, pp. 85–90.
[9] Lemmer, J. Confidence factors, empiricism and the Dempster-Shafer theory of evidence, Proc. of 1st Conf. on UAI, 1985, pp. 160–176.
[10] Voorbraak, F. On the justification of Demspster's rule of combination, Utrecht Univ., Netherlands, Logic Group Preprint Series, No. 42, 1988.
[11] Wang, P. A defect in Dempster-Shafer theory, in Proc. of 10th Conf. on Uncertainty in AI, 1994, pp. 560–566.
[12] Pearl, J. Reasoning with belief functions: An analysis of compatibility, Int. Journal of Approximate Reasoning, Vol. 4, 1990, pp. 363–389.
[13] Walley, P. Statistical Reasoning with Imprecise Probabilities, Chapman and Hall, London, 1991, pp. 278–281.
[14] Gelman, A. The boxer, the wrestler, and the coin flip: a paradox of robust Bayesian inference and belief functions, American Statistician, Vol. 60, No. 2, 2006, pp. 146–150.
[15] Smarandache F., Dezert J. Advances and applications of DSmT for information fusion, Volumes 1, 2 & 3, ARP, 2004–2009 (http://www.gallup.unm.edu/~smarandache/DSmT.htm).
[16] Smets P., Kennes R. The transferable belief model, Artif. Int., Vol. 66, 1994, pp. 191–234.
[17] Smets, P. The transferable belief model for quantified belief representation, Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol.1, Kluwer, 1998.
[18] Dezert J., Tchamova A. On the behavior of Dempster's rule of combination, School on Belief Functions Theory and Applications, Autrans, France, 4–8 April 2011 (http://hal.archives-ouvertes.fr/hal-00577983/).
[19] Dezert J., Tchamova A., Dambreville F. On the mathematical theory of evidence and Dempster's rule of combination, May 2011 (http://hal.archives-ouvertes.fr/hal-00591633/fr/).
[20] Dezert J., Wang P., Tchamova A. On The Validity of Dempster-Shafer Theory, Proc. of 15th Int. Conf. of Information Fusion, July 9-12, 2012.

Comparative Study of Contradiction Measures in the Theory of Belief Functions

Florentin Smarandache
Deqiang Han
Arnaud Martin

Originally published as: Smarandache F., Han D., Martin A. - *Comparative Study of Contradiction Measures in the Theory of Belief Functions*, in Proceedings of the 15th International Conference on Information Fusion, Singapore, 9-12 July 2012, and reprinted with permission.

***Abstract*—Uncertainty measures in the theory of belief functions are important for the uncertainty representation and reasoning. Many measures of uncertainty in the theory of belief functions have been introduced. The degree of discord (or conflict) inside a body of evidence is an important index for measuring uncertainty degree. Recently, distance of evidence is used to define a contradiction measure for quantifying the degree of discord inside a body of evidence. The contradiction measure is actually the weighted summation of the distance values between a given basic belief assignment (bba) and the categorical bba's defined on each focal element of the given bba redefined in this paper. It has normalized value and can well characterize the self-discord incorporated in bodies of evidence. We propose here, some numerical examples with comparisons among different uncertainty measures are provided, together with related analyses, to show the rationality of the proposed contradiction measure.**

***Index Terms*—Evidence theory, uncertainty measure, belief function, discord, conflict.**

I. Introduction

Dempster-Shafer evidence theory [1], also known as theory of belief functions, is one of the important uncertainty reasoning tools. It has been widely used in many applications. Evidence theory can be seen as a generalization of probability theory, where the additivity axiom is excluded. In probability theory, Shannon entropy [2] is often used for quantifying uncertainty while in the framework of evidence theory, there also need the uncertainty measure for quantifying the degree of uncertainty incorporated in a body of evidence (BOE).

In uncertainty theories, we can consider two types of uncertainty including discord (or conflict) and non-specificity, hence ambiguity [3]. There have emerged several types of uncertainty measures in the theory of belief functions. They are either the generalization of Shannon entropy and other types of uncertainty measures in probability theory or are established based on the conflict obtained by using some combination rule. For example, non-specificity [4] proposed by Dubois and Prade is a generalization of Hartley entropy [5]; aggregate uncertainty (AU) measure [6] and ambiguity measure (AM) [3] can be regarded as the generalized forms of Shannon entropy. In Martin's work [7], [8], the auto-conflict measure was proposed based on the conjunctive combination rule. There are also lots of other types of uncertainty measures in the theory of belief functions (See details in [3], [9], [11]). All the available uncertainty measures characterize the uncertainty either from one aspect (*e.g.* non-specificity and discord) or as a whole, *i.e.* the total uncertainty (*e.g.*, AM and AU).

Like in [7], [11], we attempt to break the traditional ways to establish uncertainty measure in the theory of belief functions. That is, we do not generalize the uncertainty measures in probability theory or use combination rule to obtain the uncertainty measures in theory of belief functions. In this paper we modify the contradiction measure proposed in [11] to characterize the internal conflict (or discord) degree of the uncertainty in bba's. For a bba with L focal elements, based on each focal element, a categorical bba (a bba with a unique focal element) can be obtained. Thus there are totally L categorical bba's. We calculate Jousselme's distance of evidence [10] between the original given bba and each categorical bba then we can obtained L values of distance. By using the masses of the given bba to generate the weights and executing weighted summation of the corresponding L distance values, the contradiction can be obtained. To make the contradiction measure be normalized, the normalization factor is designed and added. Some simulation results are provided to verify the correctness of the normalization factor. This contradiction measure can well characterize the conflict incorporated in a BOE, *i.e.* the self-conflict or internal conflict. Some numerical examples with comparisons among different uncertainty measures in the theory of belief functions are also provided to show the rationality of the proposed contradiction measure. It should be noted that this work is based on our previous paper [11]. The idea of constructing contradiction measure based on distance of evidence is first preliminarily proposed in that paper, where there exist some errors in the definition -corrected here- and related analyses are far from enough.

II. Basics in the theory of belief functions

A. Basic concepts in the theory of belief functions

In Dempster-Shafer evidence theory [1], The elements in the frame of discernment (FOD) (denoted by Θ) are mutually exclusive and exhaustive. Suppose that 2^Θ denotes the power-set of FOD and define the function $m : 2^\Theta \to [0, 1]$ as the

basic belief assignment (bba) satisfying:

$$\sum_{A\subseteq\Theta} m(A) = 1,\ m(\emptyset) = 0 \tag{1}$$

A bba is also called a mass function. Belief function (Bel) and plausibility function (Pl) are defined below, respectively:

$$Bel(A) = \sum_{B\subseteq A} m(B) \tag{2}$$

$$pl(A) = \sum_{A\cap B\neq\emptyset} m(B) \tag{3}$$

Suppose there are two bba's: m_1, m_2 over the FOD Θ with focal elements $A_1,\ldots,A_k$ and $B_1,\ldots,B_l$, respectively. If $k = \sum_{A_i\cap B_j=\emptyset} m_1(A_i)m_2(B_j) < 1$, $m : 2^\Theta \to [0,1]$ denoted by

$$m(A) = \begin{cases} 0,\ A = \emptyset \\ \dfrac{\sum\limits_{A_i\cap B_j=A} m_1(A_i)m_2(B_j)}{1-\sum\limits_{A_i\cap B_j=\emptyset} m_1(A_i)m_2(B_j)}, A\neq\emptyset \end{cases} \tag{4}$$

is a bba. The rule defined in Eq. (4) is called Dempster's rule of combination. In Dempster's rule of combination,

$$K = 1 - \sum_{A_i\cap B_j=\emptyset} m_1(A_i)m_2(B_j) \tag{5}$$

is used to represent the conflict between two BOEs. In recent research [12], both K and distance of evidence are used to construct a two tuple to represent the conflict between BOEs.

B. Uncertainty measures in the theory of belief functions

In the theory of belief functions, a BOE hides two types of uncertainty: non-specificity [4] and discord, hence ambiguity [3]. The available related definitions on degree of uncertainty in the theory of belief functions are briefly introduced below.

1) Auto-conflict

A n-order auto-conflict measure was proposed in [7] based on non-normalized conjunctive combination rule [13].

$$a_n = \left(\bigoplus_{i=1}^{n} m\right)(\emptyset) \tag{6}$$

The conjunctive combination rule $\oplus$ is defined as

$$m_{Conj}(C) = \sum_{A\cap B=C} m_1(A)m_2(B) := (m_1\oplus m_2)(C) \tag{7}$$

When $n = 2$, the auto-conflict equals to K in Dempster's rule of combination.

2) Non-specificity

$$N(m) = \sum_{A\subseteq\Theta} m(A)\log_2|A| \tag{8}$$

Non-specificity can be seen as weighted sum of the Hartley measure for different focal elements.

3) Confusion

Höhle proposed the measure of confusion [14] by using bba and belief function in spirit of entropy as follows.

$$Confusion(m) = -\sum_{A\in\Theta} m(A)\log_2(Bel(A)) \tag{9}$$

4) Dissonance

Yager proposed the measure of Dissonance [14] by using bba and plausibility function in spirit of entropy as follows.

$$Dissonance(m) = -\sum_{A\in\Theta} m(A)\log_2(Pl(A)) \tag{10}$$

5) Aggregate Uncertainty measure (AU)

There have emerged several definitions aiming to represent the total uncertainty in the theory of belief functions. The most representational one is a kind of generalized Shannon entropy [2], *i.e.* the aggregated uncertainty (AU) [6].

Let Bel be a belief measure on the FOD Θ. The AU associated with Bel is measured by:

$$\text{AU}(Bel) = \max_{\mathcal{P}_{Bel}}[-\sum_{\theta\in\Theta} p_\theta\log_2 p_\theta] \tag{11}$$

where the maximum is taken over all probability distributions that are consistent with the given belief function. $\mathcal{P}_{Bel}$ consists of all probability distributions $\langle p_\theta|\theta\in\Theta\rangle$ satisfying:

$$\begin{cases} p_\theta\in[0,1], \forall\theta\in\Theta \\ \sum_{\theta\in\Theta} p_\theta = 1 \\ Bel(A)\leq\sum_{\theta\in A} p_\theta\leq 1 - Bel(\bar{A}),\ \forall A\subseteq\Theta \end{cases} \tag{12}$$

As illustrated in Eq. (11) and Eq. (12), in the definition of AU, the calculation of AU is an optimization problem and bba's (or belief functions) are used to establish the constraints of the optimization problem. It is also called the "upper entropy". AU is an aggregated total uncertainty (ATU) measure, which can capture both non-specificity and discord.

AU satisfies all the requirements for uncertainty measure [9], which include probability consistency, set consistency, value range, sub-additivity and additivity for the joint BPA in Cartesian space. However, AU has the following shortcomings [3]: high computing complexity, high insensitivity to the changes of evidence, etc.

6) Ambiguity Measure (AM)

Jousselme *et al* [3] proposed AM (ambiguity measure) aiming to describe the non-specificity and discord in the theory of belief functions. Let $\Theta = \{\theta_1,\theta_2,\ldots,\theta_n\}$ be a FOD. Let m be a bba defined on Θ. Define

$$\text{AM}(m) = -\sum_{\theta\in\Theta} \text{BetP}_m(\theta)\log_2(\text{BetP}_m(\theta)) \tag{13}$$

where $\text{BetP}_m(\theta) = \sum_{\theta\in B, B\subseteq\Theta} m(B)/|B|$ is the pignistic probability distribution proposed by Smets [16]. Jousselme *et al* [3] declared that the ambiguity measure satisfies the requirements of uncertainty measure and at the same time it overcomes the defects of AU, but in fact AM does not satisfy the sub-additivity which has been pointed out by Klir [17]. Moreover in the work of Abellan [9], AM has been proved to be logically non-monotonic under some circumstances.

There are also other existing uncertainty measures in the theory of belief functions, see details in related reference [3].

III. Contradiction measure based on distance of evidence

As we can see in the previous section, all the available uncertainty measures in the theory of belief functions are direct or indirect generalization of entropy defined in probability theory or are defined by using some combination rule. Hence in [11], we break such ways in spirit of entropy in probability theory. Distance of evidence is used to construct the uncertainty degree, which is called contradiction and shown below.

$$Contr_m(m) = \sum_{X\in\mathcal{X}} m(X)\cdot d(m, m_X) \tag{14}$$

where $\mathcal{X}$ represents the set of all the focal elements of $m(\cdot)$. But it should be noted that the definition in Eq. (14) is not a normalized value. We should obtain a normalized definition for the convenience of use.

The maximum contradiction measure for $m(\cdot)$ defined on $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$ occurs when $m(\cdot)$ has a uniform distribution:

$$m(\{\theta_1\}) = m(\{\theta_2\}) = \cdots = m(\{\theta_n\}) = \frac{1}{n}$$

It depends on the cardinality of Θ and the distance used.

For $|\Theta| = n$, we use Jousselme's distance, we get max $Contr_m = \sqrt{\frac{n-1}{2n}}$.

Proof:

$$Contr_m = n\cdot\frac{1}{n}\cdot d(m, m_{\theta_i}) = d(m, m_{\theta_i})$$

i.e.: where

$$\begin{cases} m_{\theta_i}(\{\theta_i\}) = 1, \\ m_{\theta_i}(\{\theta_j\}) = 0, j\neq i, j = 1, ..., n \end{cases}$$

But the distance between m and m_{θ_i} is the same,

$$d(m, m_{\theta_i}) = \sqrt{(m-m_{\theta_i})^T \mathbf{Jac}(m-m_{\theta_i})}$$

$$= \sqrt{0.5\left[\tfrac{n-1}{n}, -\tfrac{1}{n}, \cdots, -\tfrac{1}{n}\right]\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}\begin{bmatrix} \frac{n-1}{n} \\ -\frac{1}{n} \\ \vdots \\ -\frac{1}{n} \end{bmatrix}}$$

$$= \sqrt{0.5\left[\tfrac{n-1}{n}, -\tfrac{1}{n}, \cdots, -\tfrac{1}{n}\right]\begin{bmatrix} \frac{n-1}{n} \\ -\frac{1}{n} \\ \vdots \\ -\frac{1}{n} \end{bmatrix}}$$

$$= \sqrt{0.5\left[\left(\tfrac{n-1}{n}\right)^2 + (n-1)\cdot\tfrac{1}{n^2}\right]}$$

$$= \sqrt{0.5\tfrac{(n-1)^2+n-1}{n^2}} = \sqrt{0.5\tfrac{n^2-n}{n^2}} = \sqrt{\tfrac{n-1}{2n}}$$

Therefore, in this paper, we use the normalized factor $\sqrt{\frac{n-1}{2n}}$ and then the correct normalized contradiction measure is defined below:

$$Contr_m(m) = \sqrt{\frac{2n}{n-1}}\cdot\sum_{X\in\mathcal{X}} m(X)\cdot d(m, m_X) \tag{15}$$

To further verify the correctness of the normalization factor, we design the experiments as follows.

Randomly generate 500 bba's and calculate their corresponding contradiction values based on Eq. (15). The method to randomly generate bba's is as follows [18].

Input: Θ : Frame of discernment;
N_{max}: Maximum number of focal elements
Output: Bel: Belief function (under the form of a bba, m)
Generate the power set of Θ $\mathcal{P}(\Theta)$;
Generate a random permutation of $\mathcal{P}(\Theta) \to \mathcal{R}(\Theta)$;
Generate a integer between 1 and $N_{max} \to k$;
FOReach First k elements of $\mathcal{R}(\Theta)$ do
Generate a value within $[0, 1] \to m_i, i = 1, ..., k$;
END
Normalize the vector $m = [m_1, ..., m_k] \to m'$;
$m(A_k) = m_k$;
Algorithm 1: Random generation of bba

Based on the above algorithm, the bba's generated have random number of focal elements. We set the cardinality of FOD to be 3 and 4, respectively in each experiment. Thus we totally do two experiments and the experimental results are illustrated in Fig.1.

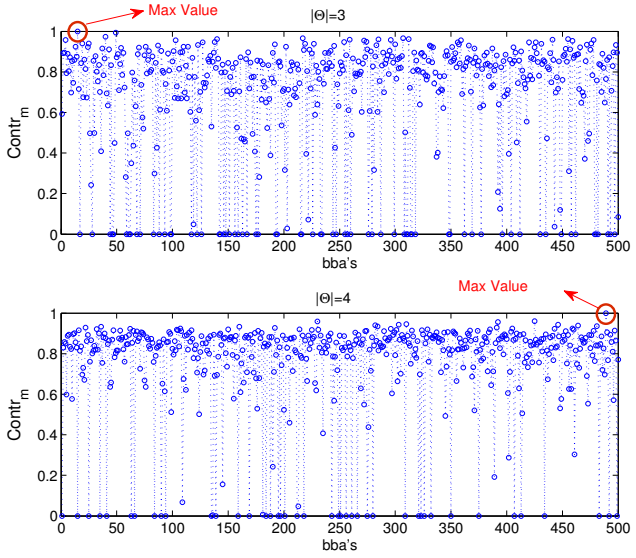


Fig. 1. Values of contradiction $Contr_m$

As shown in Fig.1, when $|\Theta| = 3$, the max value (one) is obtained at the 15th bba, which is:

$$m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = 1/3.$$

When $|\Theta| = 4$, the max value (one) is obtained at the 489th bba, which is:

$$m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = m(\{\theta_4\}) = 1/4.$$

From the proof and the experiments above, it can be seen that the selection of normalized factor is correct.

IV. Examples

A. Example 1

In this experiment, we use the bba's with focal elements of singletons and the total set. Suppose that the FOD is $\Theta = \{\theta_1, \theta_2, ..., \theta_5\}$. The initial bba is

$$m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = m(\{\theta_4\}) = m(\{\theta_5\}) = 0;$$
$$m(\Theta) = 1$$

Then at each step, the mass of $m(\Theta)$ decreases by $\Delta = 0.05$, and the mass of each $m(\{\theta_i\})$ increase by $\Delta/5 = 0.01$, where $i = 1, ..., 5$. After 20 steps, $m(\Theta)$ will become zero and $m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = m(\{\theta_4\}) = m(\{\theta_5\}) = 0.2$. Then the experiment will finish.

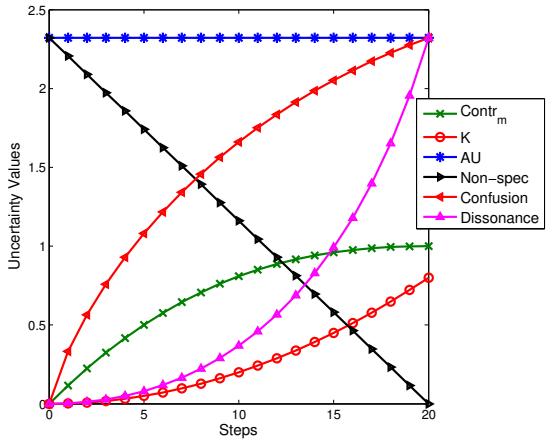


Fig. 2. Comparisons among different uncertainty measures in Example 1

As we can see in Fig. 2, although AU is deemed as a total uncertainty measure, we cannot detect the change of bba in each step based on AU.

The values of non-specificity decrease with the increase of masses of singletons.

For contradiction, K, dissonance and confusion, their values all increase with the increase of masses of singletons. Contradiction increases faster than K in the first half of all the steps and then it increases slower than K in the second half. Confusion increases faster than dissonance in the first half of all the steps and it increases slower than dissonance in the second half. The change trends of contradiction and confusion are more rational. Because at the first half of all the steps, the relative changes of the masses of singletons increase more significantly than the relative changes in the second half.

The value of contradiction belongs to $[0, 1]$ and it reaches its maximum value at the final step, *i.e.*:

When $m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = m(\{\theta_4\}) = m(\{\theta_5\}) = 0.2$, $Contr_m = 1$

B. Example 2

In this experiment, we use the bba's with focal elements of singletons and the total set. Suppose that the FOD is $\Theta = \{\theta_1, \theta_2, ..., \theta_5\}$. The initial bba is

$$m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = m(\{\theta_4\}) = m(\{\theta_5\}) = 0;$$
$$m(\Theta) = 1$$

Then at each step, the mass of $m(\Theta)$ decreases by $\Delta = 0.05$, and the mass of one singleton $m(\{\theta_1\})$ increase by $\Delta = 0.05$ at each step. After 20 steps, $m(\Theta)$ will become zero and the experiment will finish.

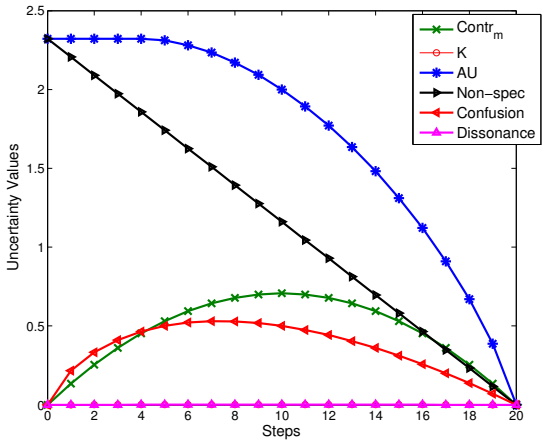


Fig. 3. Comparisons among different uncertainty measures in Example 2

As we can see in Fig. 3, with the increase of $m(\{\theta_1\})$ and the decrease of $m(\Theta)$ in each step, the AU and non-specificity decrease.

Although for the original bba, the non-specificity is highest, the conflict inside should be the least. So AU can not characterize the discord part of the uncertainty incorporated in the BOE.

K and Dissonance cannot detect the change of bba.

The value of the proposed contradiction increases at first and reaches the max value when the bba becomes

$$m(\{\theta_1\}) = 0.5, m(\Theta) = 0.5$$

Then with the increase of $m(\{\theta_1\})$ and the decrease of $m(\Theta)$ in following steps, the value of the proposed contradiction decrease and it reach zero when $m(\{\theta_1\}) = 1$, which is the clearest case.

If we consider the two focal elements $\{\theta_1\}$ and Θ are different in the power-set of Θ, when their values are equal the uncertainty reaches the max value. This should be more rational.

Confusion has the similar change trend compared to that of our proposed contradiction measure. But the maximum value of confusion does not occur at the middle.

C. Example 3

In this experiment, we use the bba's with focal elements of the same cardinality. Suppose that the FOD is $\Theta = \{\theta_1, \theta_2, ..., \theta_5\}$.

The initial bba is

$$m(\{\theta_1,\theta_2\}) = m(\{\theta_1,\theta_3\}) = m(\{\theta_1,\theta_4\})$$
$$= m(\{\theta_2,\theta_3\}) = m(\{\theta_2,\theta_4\}) = 0; m(\{\theta_3,\theta_4\}) = 1$$

Then at each step, the mass of $m(\{\theta_3,\theta_4\})$ decreases by $\Delta = 0.05$, and the masses of all the other focal elements increase by $\Delta = 0.05/5 = 0.01$ at each step. After 16 steps, masses of all the focal elements become equal.

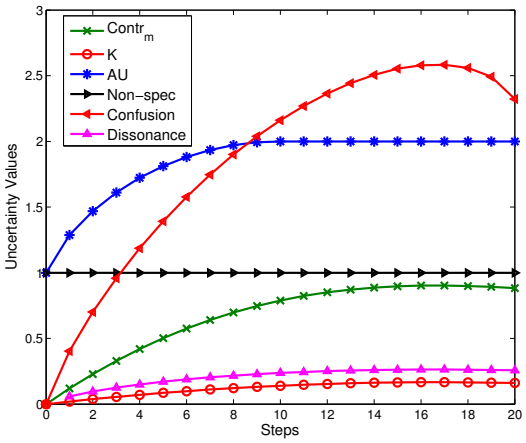


Fig. 4. Comparisons among different uncertainty measures in Example 3

As we can see in Fig. 4, Non-specificity can not detect the change of bba. This is because Non-specificity mainly concerns the cardinality of focal elements.

AU can detect the change of bba, but after step 10, the values of AU are the same with the change of bba in following steps. Thus AU is not sensitive to the change of bba.

With the change of bba in each step, K and Dissonance change very little. Thus here K and dissonance are not so sensitive to the change of bba.

For contradiction proposed and confusion, they can detect the change of bba well.

D. Example 4

Suppose that the FOD is $\Theta = \{\theta_1,\theta_2\}$. The initial bba is

$$m(\{\theta_1\}) = a, \quad m(\{\theta_2\}) = b,$$
$$m(\{\theta_1,\theta_1\}) = 1 - a - b.$$

Suppose that $a, b \in [0, 0.5]$, we calculate the values of all the uncertainty measures according to the change of a and b

As we can see in Fig. 5, with the change of a and b, AU are always the same.

All the other measures can detect the change of a and b.

We can see that the value of the proposed contradiction varies relatively uniformly when compared with other meaures. Thus the contradiction is not too sensitive and at the same time not too insensitive to the change of bba.

The value range belongs to [0,1], which is good characteristic for being a measure for quantifying the degree of discord.

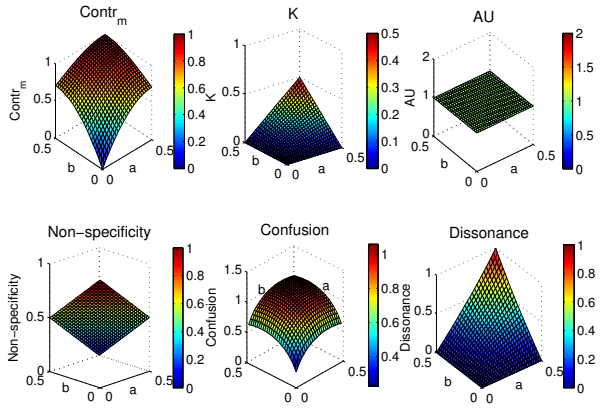


Fig. 5. Comparisons among different uncertainty measures in Example 4

V. Further Analysis

In definition of $Contr_m$ in Eq. (15), the distance used is Jousselme's distance. In our work, we have also tried other types of distances in the theory of belief functions to construct the contradiction, which include

1) Betting commitment distance (Pignistic probability distance)

$$d_T(m_1,m_2) = \max_{A\subseteq\Theta}\{|\mathrm{BetP}_1(A) - \mathrm{BetP}_2(A)|\} \tag{16}$$

where BetP represents the pignistic probability of corresponding bba.

2) Cuzzonlin distance

$$d_{Cuzz}(m_1,m_2) = \sqrt{(m_1,m_2)^T \mathbf{IncInc}^T (m_1,m_2)} \tag{17}$$

where **Inc** is

$$\begin{cases} \mathbf{Inc}(A,B) = 1, if\, A \subseteq B \\ 0, \text{others} \end{cases} \tag{18}$$

3) Conflict distance

$$d_K((m_1,m_2)) = m_1^T(\mathbf{I} - \mathbf{Inc})m_2 \tag{19}$$

4) Bhattacharyya distance

$$d_B(m_1,m_2) = (1 - \sqrt{m_1}^T \mathbf{I} \sqrt{m_2})^p \tag{20}$$

We do following experiments to compare the different contradiction measures defined on the different distance definitions above. When we use d_{Cuzz} and d_K to construct normalized contradiction measures, the normalization factor should be $(n-1)/n$.

A. Example 5

Suppose that the FOD is $\Theta = \{\theta_1,\theta_2,\theta_3\}$.

The initial bba is

$$m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = 0; m(\Theta) = 1$$

Then at each step, the mass of $m(\Theta)$ decreases by $\Delta = 0.05$, and the mass of each $m(\{\theta_i\})$ increase by $\Delta/3 = 0.05/3$,where $i = 1, 2, 3$.

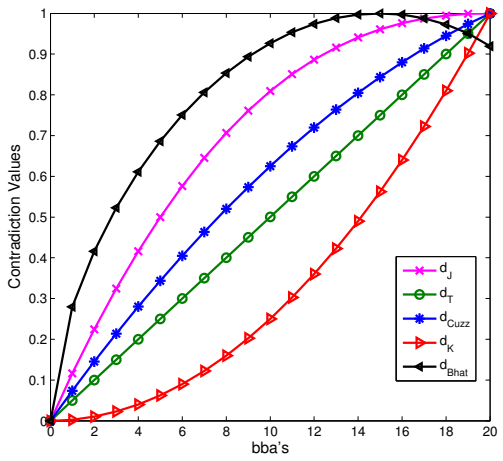


Fig. 6. Comparisons among different contradiction measures based on different distance measures - Example 5

B. Example 6

Suppose that the FOD is $\Theta = \{\theta_1, \theta_2, \theta_3\}$.
The initial bba is

$$m(\{\theta_1\}) = m(\{\theta_2\}) = m(\{\theta_3\}) = 0; m(\Theta) = 1$$

Then at each step, the mass of $m(\Theta)$ decreases by $\Delta = 0.05$, and the mass of $m(\{\theta_1\})$ increase by $\Delta = 0.05$. In the final step, the bba obtained is

$$m(\{\theta_1\}) = 1$$
$$m(\{\theta_2\}) = m(\{\theta_3\}) = m(\Theta) = 0;$$

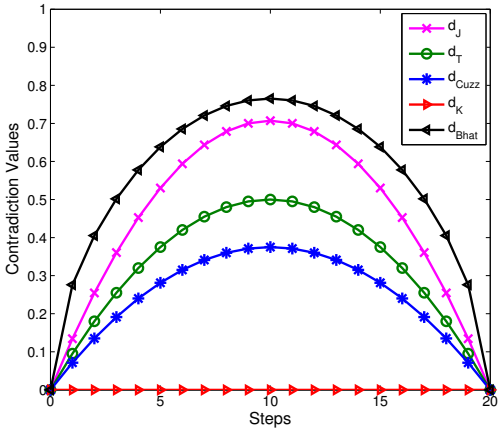


Fig. 7. Comparisons among different contradiction measures based on different distance measures - Example 5

As we can see in Example 5 and 6, all the contradiction measures obtained based on different distance definitions can well characterize the degree of discord inside BOEs. Till now, only Jousselme's distance is a strict distance metric, so we suggest to use Jousselme's distance.

VI. Conclusion

In this paper, we propose a new normalization of a measure called contradiction to characterize the degree of discord or conflict inside a body of evidence. This contradiction measure is distance-based and it can well describe the discord part of the uncertainty in the theory of belief functions. Some numerical examples are provided to support the rationality of the proposed contradiction measure.

In our work, we have also preliminarily tried other types of distance in evidence theory to construct the contradiction measure. In our future work, we will further analyze the contradiction defined on different distance measures. Contradiction measure can represent the qualities of different information sources to some extent. Thus we will also try to use the contradiction measure in applications based on the evaluation of bba's, for example, the weights determination in weighted evidence combination.

References

[1] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University, Princeton, 1976.
[2] C.E. Shannon, "A mathematical theory of communication", *Bell System Technical Journal* vol. 27, pp. 379–423, 623–656, 1948.
[3] A.L. Jousselme, C.S. Liu, D. Grenier, E. Bosse, "Measuring ambiguity in the evidence theory", *IEEE Transactions on Systems, Man and Cybernetics, Part A*, vol. 36, no. 5, pp. 890–903, 2006.
[4] D. Dubois, H. Prade, "A note on measures of specificity for fuzzy sets", *International Journal of General Systems*, vol. 10, no. 4, pp. 279–283, 1985.
[5] R.V.L. Hartley, "Transmission of information", *Bell System Technical Journal* vol. 7, pp. 535–563, 1928.
[6] D. Harmanec, G.J. Klir, "Measuring total uncertainty in Dempster-Shafer theory", *International Journal of General Systems*, vol. 22, no. 4, pp. 405–419, 1994.
[7] A. Martin, A.-L. Jousselme and C. Osswald, "Conflict measure for the discounting operation on belief functions," Proc. of International Conference on Information Fusion, Cologne, Germany, 30 June-3 July 2008.
[8] C. Osswald and A. Martin, "Understanding the large family of Dempster-Shafer theory's fusion operators - a decision-based measure", *International Conference on Information Fusion*, Florence, Italy, 10-13 July 2006.
[9] J. Abellan and A. Masegosa. "Requirements For Total Uncertainty Measures In Dempster-shafer theory Of Evidence", *International Journal of General Systems*, vol. 37, no. 6, pp. 733–747, 2008.
[10] A.-L. Jousselme, D. Grenier, E. Bosse, "A new distance between two bodies of evidence," *Information Fusion*, vol. 2, no. 2, pp. 91–101, June 2001.
[11] F. Smarandache, A. Martin, C. Osswald, "Contradiction measures and specificity degrees of basic belief assignments,". *Proc. of 14th International Conference on Information Fusion*, Chicago, USA, 2011, pp. 475-482.
[12] W.R. Liu. "Analyzing the degree of conflict among belief functions," *Artificial Intelligence*, vol. 170, no. 11, pp. 909 - 924, 2006.
[13] P. Smets, "The Combination of Evidence in the Transferable Belief Model, " *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 447-458, 1990
[14] U. Höhle, "Entropy with respect to plausibility measures". *Proceedings of 12th IEEE International Symposium on Multiple-Valued Logic*, 1982, pp. 167-169.
[15] R. R. Yager, "Entropy and specificity in a mathematical theory of evidence". *International Journal of General Systems*, vol. 9, no. 4, pp. 249-260, 1983.
[16] P. Smets, R. Kennes, "The transferable belief model", *Artificial Intelligence*, vol. 66, no. 2, pp. 191-234, 1994.
[17] G. J. Klir and H. W. Lewis, Remarks on "Measuring Ambiguity in the Evidence Theory", *IEEE Transactions on Systems, Man and Cybernetics, Part A*, vol. 38, no. 4, pp. 995–999, 2008.
[18] A.-L. Jousselme, P. Maupin, "On some properties of distances in evidence theory," in *Workshop on Theory of Belief Functions*, Brest, France, Mar 31st - Apr 2nd, 2010, pp. 1-6.

Why Dempster's Rule doesn't behave as Bayes rule with Informative Priors

Jean Dezert
Albena Tchamova
Deqiang Han
Jean-Marc Tacnet

Originally published as Dezert J., Tchamova A., Han D., Tacnet J.-M., *Why Dempster's rule doesn't behave as Bayes rule with informative priors,* Proc. of 2013 IEEE International Symposium on Innovations in Intelligent Systems and Application (INISTA 2013), Albena, Bulgaria, June 19-21, 2013, and reprinted with permission.

***Abstract*—In this paper, we analyze Bayes fusion rule in details from a fusion standpoint, as well as the emblematic Dempster's rule of combination introduced by Shafer in his Mathematical Theory of evidence based on belief functions. We propose a new interesting formulation of Bayes rule and point out some of its properties. A deep analysis of the compatibility of Dempster's fusion rule with Bayes fusion rule is done. Our analysis proves clearly that Dempster's rule of combination does not behave as Bayes fusion rule in general, because these methods deal very differently with the prior information when it is really informative (not uniform). Only in the very particular case where the basic belief assignments to combine are Bayesian and when the prior information is uniform (or vacuous), Dempster's rule remains consistent with Bayes fusion rule. In more general cases, Dempster's rule is incompatible with Bayes rule and it is not a generalization of Bayes fusion rule.**

***Keywords*—Information fusion, Probability theory, Bayes fusion rule, Dempster's fusion rule.**

I. Introduction

In 1979, Lotfi Zadeh questioned in [1] the validity of the Dempster's rule of combination [2], [3] proposed by Shafer in Dempster-Shafer Theory (DST) of evidence [4]. Since more than 30 years many strong debates [5], [6], [7], [8], [9], [10], [11], [12], [13] on the validity of foundations of DST and Dempster's rule have bloomed. The purpose of this paper is not to discuss the validity of Dempster's rule, nor the foundations of DST which have been already addressed in previous papers [14], [15], [16]. In this paper, we just focus on the deep analysis of the real incompatibility of Dempster's rule with Bayes fusion rule. Our analysis supports Mahler's one briefly presented in [17]. This paper is organized as follows. In section II, we recall basics of conditional probabilities and Bayes fusion rule with its main properties. In section III, we recall the basics of belief functions and Dempster's rule. In section IV, we analyze in details the incompatibility of Dempster's rule with Bayes rule in general and its partial compatibility for the very particular case when prior information is modeled by a Bayesian uniform basic belief assignment (bba). Section V concludes this paper.

II. Conditional probabilities and Bayes fusion

In this section, we recall the definition of conditional probability [18] and present the principle and the properties of Bayes fusion rule. We present the structure of this rule derived from the classical definition of the conditional probability in a new uncommon interesting form that will help us to analyze its partial similarity with Dempster's rule proposed by Shafer in his mathematical theory of evidence [4]. We will show clearly why Dempster's rule fails to be compatible with Bayes rule in general.

A. Conditional probabilities

Let us consider two random events X and Z. The conditional probability mass functions (pmfs) $P(X|Z)$ and $P(Z|X)$ are defined (assuming $P(X) > 0$ and $P(Z) > 0$) by [18]:

$$P(X|Z) \triangleq \frac{P(X \cap Z)}{P(Z)} \quad \text{and} \quad P(Z|X) \triangleq \frac{P(X \cap Z)}{P(X)} \tag{1}$$

which yields to Bayes Theorem:

$$P(X|Z) = \frac{P(Z|X)P(X)}{P(Z)} \text{ and } P(Z|X) = \frac{P(X|Z)P(Z)}{P(X)} \tag{2}$$

where $P(X)$ is called the *a priori probability* of X, and $P(Z|X)$ is called the *likelihood of* X. The denominator $P(Z)$ plays the role of a normalization constant.

B. Bayes parallel fusion rule

In fusion applications, we are often interested in computing the probability of an event X given two events Z_1 and Z_2 that have occurred. More precisely, one wants to compute $P(X|Z_1 \cap Z_2)$ knowing $P(X|Z_1)$ and $P(X|Z_2)$, where X can take N distinct exhaustive and exclusive states x_i, $i = 1, 2, \ldots, N$. Such type of problem is traditionally called a *fusion problem*. $P(X|Z_1 \cap Z_2)$ becomes easily computable by assuming the following conditional statistical independence condition expressed mathematically by:

$$(A1): \qquad P(Z_1 \cap Z_2|X) = P(Z_1|X)P(Z_2|X) \tag{3}$$

With such conditional independence condition (A1), then from Eq. (1) and Bayes Theorem one gets:

$$P(X|Z_1 \cap Z_2) = \frac{\frac{P(X|Z_1)P(X|Z_2)}{P(X)}}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)}} \tag{4}$$

The rule of combination given by Eq. (4) is known as *Bayes parallel (or product) rule* and dates back to Bernoulli [19]. The

Eq. (4) can be rewritten as:

$$P(X|Z_1 \cap Z_2) = \frac{1}{K(X,Z_1,Z_2)} \cdot P(X|Z_1) \cdot P(X|Z_2) \quad (5)$$

where the coefficient $K(X,Z_1,Z_2)$ is defined by:

$$K(X,Z_1,Z_2) \triangleq P(X) \cdot \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)} \quad (6)$$

C. Symmetrization of Bayes fusion rule

The expression of Bayes fusion rule given by Eq. (4) can also be symmetrized in the following form that, quite surprisingly, rarely appears in the literature:

$$P(X|Z_1 \cap Z_2) = \frac{1}{K'(Z_1,Z_2)} \cdot \frac{P(X|Z_1)}{\sqrt{P(X)}} \cdot \frac{P(X|Z_2)}{\sqrt{P(X)}} \quad (7)$$

where the normalization constant $K'(Z_1,Z_2)$ is given by:

$$K'(Z_1,Z_2) \triangleq \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}} \quad (8)$$

We call the quantity $A_2(X=x_i) \triangleq \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}}$ entering in Eq. (8) the *Agreement Factor* on $X=x_i$ of order 2. The level of the Global Agreement (GA) of the conjunctive consensus taking into account the prior pmf of X is represented as:

$$GA_2 \triangleq \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}} = K'(Z_1,Z_2) \quad (9)$$

In fact, with assumption (A1), the probability $P(X|Z_1 \cap Z_2)$ given in Eq. (7) is nothing but the simple ratio of the agreement factor $A_2(X)$ on X over the global agreement $GA_2 = \sum_{i=1}^{N} A_2(X=x_i)$, that is:

$$P(X|Z_1 \cap Z_2) = \frac{A_2(X)}{GA_2} \quad (10)$$

The quantity GC_2 measures the global conflict (i.e. the total conjunctive disagreement) taking into account the prior pmf of X.

$$GC_2 \triangleq \sum_{i_1,i_2=1|i_1 \neq i_2}^{N} \frac{P(X=x_{i_1}|Z_1)}{\sqrt{P(X=x_{i_1})}} \cdot \frac{P(X=x_{i_2}|Z_2)}{\sqrt{P(X=x_{i_2})}} \quad (11)$$

• Symbolic representation of Bayes fusion rule

The (symmetrized form of) Bayes fusion rule of two posterior probability measures $P(X|Z_1)$ and $P(X|Z_2)$, given in Eq. (7), requires an extra knowledge of the prior probability of X. For convenience, we denote symbolically this fusion rule as:

$$P(X|Z_1 \cap Z_2) = Bayes(P(X|Z_1), P(X|Z_2); P(X)) \quad (12)$$

• Particular case: Uniform a priori pmf

In such particular case, all the prior probabilities values $\sqrt{P(X=x_i)} = \sqrt{1/N}$ and $\sqrt[s]{P(X=x_i)} = \sqrt[s]{1/N}$ can be simplified in Bayes fusion formulas Eq. (7) and Eq. (8). Therefore, Bayes fusion formula (7) reduces to:

$$P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1)P(X|Z_2)}{\sum_{i=1}^{N} P(X=x_i|Z_1)P(X=x_i|Z_2)} \quad (13)$$

By convention, Eq. (13) is denoted symbolically as:

$$P(X|Z_1 \cap Z_2) = Bayes(P(X|Z_1), P(X|Z_2)) \quad (14)$$

Similarly, $Bayes(P(X|Z_1), \ldots, P(X|Z_s))$ rule defined with an uniform a priori pmf of X will be given by:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{\sum_{i=1}^{N} \prod_{k=1}^{s} P(X=x_i|Z_k)} \quad (15)$$

When $P(X)$ is uniform one has $GA_2^{unif} + GC_2^{unif} = 1$. Eq. (13) can be expressed as:

$$P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1)P(X|Z_2)}{GA_2^{unif}} = \frac{P(X|Z_1)P(X|Z_2)}{1 - GC_2^{unif}} \quad (16)$$

By a direct extension, one will have:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{GA_s^{unif}} = \frac{\prod_{k=1}^{s} P(X|Z_k)}{1 - GC_s^{unif}} \quad (17)$$

$$GA_s^{unif} = \sum_{i_1,\ldots,i_s=1|i_1=\ldots=i_s}^{N} P(X=x_{i_1}|Z_1) \ldots P(X=x_{i_s}|Z_s)$$

$$GC_s^{unif} = 1 - GA_s^{unif}$$

D. Properties of Bayes fusion rule

• (P1) : The pmf $P(X)$ is a neutral element of the Bayes fusion rule when combining only two sources.

Proof: A source is called a neutral element of a fusion rule if and only if it has no influence on the fusion result. $P(X)$ is a neutral element of Bayes rule if and only if $Bayes(P(X|Z_1), P(X); P(X)) = P(X|Z_1)$. It can be easily verified that this equality holds by replacing $P(X|Z_2)$ by $P(X)$ and $P(X=x_i|Z_2)$ by $P(X=x_i)$ (as if the conditioning term Z_2 vanishes) in Eq. (4). One can also verify that $Bayes(P(X), P(X|Z_2); P(X)) = P(X|Z_2)$, which completes the proof.

• (P2) : Bayes fusion rule is in general not idempotent.

Proof: A fusion rule is idempotent if the combination of all same inputs is equal to the inputs. To prove that Bayes rule is not idempotent it suffices to prove that: in general

$$Bayes(P(X|Z_1), P(X|Z_1); P(X)) \neq P(X|Z_1)$$

From Bayes rule (4), when $P(X|Z_2) = P(X|Z_1)$ we clearly get in general

$$\frac{1}{P(X)} \frac{P(X|Z_1)P(X|Z_1)}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_1)}{P(X=x_i)}} \neq P(X|Z_1) \quad (18)$$

but when Z_1 and Z_2 vanish, because in such case Eq. (18) reduces to $P(X)$ on its left and right sides.

• (P3) : Bayes fusion rule is in general not associative.

Proof: A fusion rule f is called associative if and only if it satisfies the associative law: $f(f(x,y),z) = f(x,f(y,z)) =$

$f(y, f(x, z)) = f(x, y, z)$ for all possible inputs x, y and z. Let us prove Bayes rule is not associative from a very simple example.

Example 1: Let us consider the simplest set of outcomes $\{x_1, x_2\}$ for X, with prior pmf:

$$P(X = x_1) = 0.2 \text{ and } P(X = x_2) = 0.8$$

and let us consider the three given sets of posterior pmfs:

$$\begin{cases} P(X = x_1|Z_1) = 0.1 \text{ and } P(X = x_2|Z_1) = 0.9 \\ P(X = x_1|Z_2) = 0.5 \text{ and } P(X = x_2|Z_2) = 0.5 \\ P(X = x_1|Z_3) = 0.6 \text{ and } P(X = x_2|Z_3) = 0.4 \end{cases}$$

One can see that even if in our example one has $f(x, f(y, z)) = f(f(x, y), z) = f(y, f(x, z))$ because $P(X|(Z_1 \cap Z_2) \cap Z_3) = P(X|Z_1 \cap (Z_2 \cap Z_3)) = P(X|Z_2 \cap (Z_1 \cap Z_3))$, the Bayes fusion rule is not associative since:

$$\begin{cases} P(X|(Z_1 \cap Z_2) \cap Z_3) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \\ P(X|Z_1 \cap (Z_2 \cap Z_3)) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \\ P(X|Z_2 \cap (Z_1 \cap Z_3)) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \end{cases}$$

• (P4) : Bayes fusion rule is associative if and only if $P(X)$ is uniform.

Proof: If $P(X)$ is uniform, Bayes fusion rule is given by Eq. (15) which can be rewritten as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{P(X|Z_1 \cap \ldots \cap Z_{s-1})P(X|Z_s)}{\sum_{i=1}^{N} P(X = x_i|Z_1 \cap \ldots \cap Z_{s-1})P(X = x_i|Z_s)}$$

Therefore when $P(X)$ is uniform, one has:

$$Bayes(P(X|Z_1), \ldots, P(X|Z_s)) = Bayes(Bayes(P(X|Z_1), \ldots, P(X|Z_{s-1})), P(X|Z_s)).$$

• (P5) : The levels of global agreement and global conflict between the sources do not matter in Bayes fusion rule.

Proof: This property seems surprising at first glance, but, since the results of Bayes fusion is nothing but the ratio of the agreement on x_i $(i = 1, 2, \ldots, N)$ over the global agreement factor, many distinct sources with different global agreements (and this with different global conflicts) can yield same Bayes fusion result. Indeed, the ratio is kept unchanged when multiplying its numerator and denominator by same non null scalar value. Consequently, the absolute levels of global agreement between the sources (and therefore of global conflict also) do not matter in Bayes fusion result. What really matters is only the proportions of relative agreement factors.

III. Belief functions and Dempster's rule

The Belief Functions (BF) have been introduced in 1976 by Glenn Shafer in his mathematical theory of evidence [4], also known as Dempster-Shafer Theory (DST) in order to reason under uncertainty and to model epistemic uncertainties. The emblematic fusion rule proposed by Shafer to combine sources of evidences characterized by their basic belief assignments (bba) is Dempster's rule that will be analyzed in details in the sequel. In the literature over the years, DST has been widely defended by its proponents in arguing that: 1) Probability measures are particular cases of Belief functions; and 2) Dempster's fusion rule is a generalization of Bayes fusion rule. Although the statement 1) is correct because Probability measures are indeed particular (additive) Belief functions (called as Bayesian belief functions), we will explain why the second statement about Dempster's rule is incorrect in general.

A. Belief functions

Let Θ be a frame of discernment of a problem under consideration. More precisely, the set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ consists of a list of N exhaustive and exclusive elements θ_i, $i = 1, 2, \ldots, N$. Each θ_i represents a possible state related to the problem we want to solve. The exhaustivity and exclusivity of elements of Θ is referred as Shafer's model of the frame Θ. A basic belief assignment (bba), also called a belief mass function, $m(.) : 2^\Theta \to [0, 1]$ is a mapping from the power set of Θ (i.e. the set of subsets of Θ), denoted 2^Θ, to $[0, 1]$, that verifies the following conditions [4]:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{X \in 2^\Theta} m(X) = 1 \tag{19}$$

The quantity $m(X)$ represents the mass of belief exactly committed to X. An element $X \in 2^\Theta$ is called a focal element if and only if $m(X) > 0$. The set $\mathcal{F}(m) \triangleq \{X \in 2^\Theta | m(X) > 0\}$ of all focal elements of a bba $m(.)$ is called the core of the bba. A bba $m(.)$ is said Bayesian if its focal elements are singletons of 2^Θ. The vacuous bba characterizing the total ignorance denoted $I_t = \theta_1 \cup \theta_2 \cup \ldots \cup \theta_N$ is defined by $m_v(.) : 2^\Theta \to [0; 1]$ such that $m_v(X) = 0$ if $X \neq \Theta$, and $m_v(I_t) = 1$.

From any bba $m(.)$, the belief function $Bel(.)$ and the plausibility function $Pl(.)$ are defined for $\forall X \in 2^\Theta$ as:

$$\begin{cases} Bel(X) = \sum_{Y \in 2^\Theta | Y \subseteq X} m(Y) \\ Pl(X) = \sum_{Y \in 2^\Theta | X \cap Y \neq \emptyset} m(Y) \end{cases} \tag{20}$$

$Bel(X)$ represents the whole mass of belief that comes from all subsets of Θ included in X. It is interpreted as the lower bound of the probability of X, i.e. $P_{\min}(X)$. $Bel(.)$ is a subadditive measure since $\sum_{\theta_i \in \Theta} Bel(\theta_i) \leq 1$. $Pl(X)$ represents the whole mass of belief that comes from all subsets of Θ compatible with X (i.e., those intersecting X). $Pl(X)$ is interpreted as the upper bound of the probability of X, i.e. $P_{\max}(X)$. $Pl(.)$ is a superadditive measure since $\sum_{\theta_i \in \Theta} Pl(\theta_i) \geq 1$. $Bel(X)$ and $Pl(X)$ are classically seen [4] as lower and upper bounds of an unknown probability $P(.)$, and one has the following inequality satisfied $\forall X \in 2^\Theta$: $Bel(X) \leq P(X) \leq Pl(X)$. The belief function $Bel(.)$ (and the plausibility function $Pl(.)$) built from any Bayesian bba $m(.)$ can be interpreted as a (subjective) conditional probability measure provided by a given source of evidence, because if the bba $m(.)$ is *Bayesian* the following equality always holds [4]: $Bel(X) = Pl(X) = P(X)$.

B. Dempster's rule of combination

Dempster's rule of combination, denoted DS rule is a mathematical operation, represented symbolically by $\oplus$, which corresponds to the normalized conjunctive fusion rule. Based on Shafer's model of Θ, the combination of $s > 1$ independent and distinct sources of evidences characterized by their bba

$m_1(.)$, ..., $m_s(.)$ related to the same frame of discernment Θ is denoted $m_{DS}(.) = [m_1 \oplus \ldots \oplus m_s](.)$. The quantity $m_{DS}(.)$ is defined mathematically as follows: $m_{DS}(\emptyset) \triangleq 0$ and $\forall X \neq \emptyset \in 2^\Theta$

$$m_{DS}(X) \triangleq \frac{m_{12\ldots s}(X)}{1 - K_{12\ldots s}} \tag{21}$$

where the conjunctive agreement on X is given by:

$$m_{12\ldots s}(X) \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = X}} m_1(X_1) m_2(X_2) \ldots m_s(X_s) \tag{22}$$

and where the global conflict is given by:

$$K_{12\ldots s} \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = \emptyset}} m_1(X_1) m_2(X_2) \ldots m_s(X_s) \tag{23}$$

When $K_{12\ldots s} = 1$, the s sources are in total conflict and their combination cannot be computed with DS rule because Eq. (21) is mathematically not defined due to 0/0 indeterminacy [4]. DS rule is commutative and associative which makes it very attractive from engineering implementation standpoint. It has been proved in [4] that the vacuous bba $m_v(.)$ is a neutral element for DS rule because $[m \oplus m_v](.) = [m_v \oplus m](.) = m(.)$ for any bba $m(.)$ defined on 2^Θ.

IV. Analysis of compatibility of Dempster's rule with Bayes rule

To analyze the compatibility of Dempster's rule with Bayes rule, we need to work in the probabilistic framework because Bayes fusion rule has been developed only in this theoretical framework. So in the sequel, we will manipulate only probability mass functions (pmfs), related with Bayesian bba's in the Belief Function framework. If Dempster's rule is a true (consistent) generalization of Bayes fusion rule, it must provide same results as Bayes rule when combining Bayesian bba's, otherwise Dempster's rule cannot be fairly claimed to be a generalization of Bayes fusion rule. In this section, we analyze the real (partial or total) compatibility of Dempster's rule with Bayes fusion rule. Two important cases must be analyzed depending on the nature of the prior information $P(X)$ one has in hands for performing the fusion of the sources. These sources to combine will be characterized by the following Bayesian bba's:

$$\begin{cases} m_1(.) \triangleq \{m_1(\theta_i) = P(X = x_i|Z_1), i = 1, 2, \ldots, N\} \\ \quad \vdots \quad \vdots \qquad\qquad \vdots \\ m_s(.) \triangleq \{m_s(\theta_i) = P(X = x_i|Z_s), i = 1, 2, \ldots, N\} \end{cases} \tag{24}$$

The prior information is characterized by a given bba denoted by $m_0(.)$ that can be defined either on 2^Θ, or only on Θ if we want to deal for the needs of our analysis with a Bayesian prior. In the latter case, if $m_0(.) \triangleq \{m_0(\theta_i) = P(X = x_i), i = 1, 2, \ldots, N\}$ then $m_0(.)$ plays the same role as the prior pmf $P(X)$ in the probabilistic framework.

When considering a non vacuous prior $m_0(.) \neq m_v(.)$, we denote Dempster's combination of s sources symbolically as:

$$m_{DS}(.) = DS(m_1(.), \ldots, m_s(.); m_0(.))$$

When the prior bba is vacuous $m_0(.) = m_v(.)$ then $m_0(.)$ has no impact on Dempster's fusion result, and so we denote symbolically Dempster's rule as:

$$m_{DS}(.) = DS(m_1(.), \ldots, m_s(.); m_v(.)) = DS(m_1(.), \ldots, m_s(.))$$

A. Case 1: Uniform Bayesian prior

It is important to note that Dempster's fusion formula proposed by Shafer in [4] and recalled in Eq. (21) makes no real distinction between the nature of sources to combine (if they are posterior or prior information). In fact, the formula (21) reduces exactly to Bayes rule given in Eq. (17) if the bba's to combine are Bayesian and if the prior information is either uniform or vacuous. Stated otherwise the following functional equality holds:

$$DS(m_1(.), \ldots, m_s(.); m_0(.)) \equiv Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X)) \tag{25}$$

as soon as all bba's $m_i(.)$, $i = 1, 2, \ldots, s$ are Bayesian and coincide with $P(X|Z_i)$, $P(X)$ is uniform, and either the prior bba $m_0(.)$ is vacuous ($m_0(.) = m_v(.)$), or $m_0(.)$ is the uniform Bayesian bba.

Example 2: Let us consider $\Theta(X) = \{x_1, x_2, x_3\}$ with two distinct sources providing the following Bayesian bba's:

$$\begin{cases} m_1(x_1) = P(X = x_1|Z_1) = 0.2 \\ m_1(x_2) = P(X = x_2|Z_1) = 0.3 \\ m_1(x_3) = P(X = x_3|Z_1) = 0.5 \end{cases} \quad \text{and} \quad \begin{cases} m_2(x_1) = 0.5 \\ m_2(x_2) = 0.1 \\ m_2(x_3) = 0.4 \end{cases}$$

• If we choose as prior $m_0(.)$ the vacuous bba, that is $m_0(x_1 \cup x_2 \cup x_3) = 1$, then one will get (with $K_{12}^{vacuous} = 0.67$):

$$\begin{cases} m_{DS}(x_1) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_1) m_2(x_1) m_0(x_1 \cup x_2 \cup x_3) \\ &= \frac{1}{1-0.67} 0.2 \cdot 0.5 \cdot 1 = \frac{0.10}{0.33} \approx 0.3030 \\ m_{DS}(x_2) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_2) m_2(x_2) m_0(x_1 \cup x_2 \cup x_3) \\ &= \frac{1}{1-0.67} 0.3 \cdot 0.1 \cdot 1 = \frac{0.03}{0.33} \approx 0.0909 \\ m_{DS}(x_3) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_3) m_2(x_3) m_0(x_1 \cup x_2 \cup x_3) \\ &= \frac{1}{1-0.67} 0.5 \cdot 0.4 \cdot 1 = \frac{0.20}{0.33} \approx 0.6061 \end{cases}$$

• If we choose as prior $m_0(.)$ the uniform Bayesian bba given by $m_0(x_1) = m_0(x_2) = m_0(x_3) = 1/3$, then we get:

$$\begin{cases} m_{DS}(x_1) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_1) m_2(x_1) m_0(x_1) \\ &= \frac{1}{1-0.89} 0.2 \cdot 0.5 \cdot 1/3 = \frac{0.10/3}{0.11} \approx 0.3030 \\ m_{DS}(x_2) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_2) m_2(x_2) m_0(x_2) \\ &= \frac{1}{1-0.89} 0.3 \cdot 0.1 \cdot 1/3 = \frac{0.03/3}{0.11} \approx 0.0909 \\ m_{DS}(x_3) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_3) m_2(x_3) m_0(x_3) \\ &= \frac{1}{1-0.89} 0.5 \cdot 0.4 \cdot 1/3 = \frac{0.20/3}{0.11} \approx 0.6061 \end{cases}$$

where the degree of conflict when $m_0(.)$ is Bayesian and uniform is now given by $K_{12}^{uniform} = 0.89$.

Clearly $K_{12}^{uniform} \neq K_{12}^{vacuous}$, but the fusion results obtained with two distinct priors $m_0(.)$ (vacuous or uniform) are the same because of the algebraic simplification by 1/3 in Dempster's fusion formula when using uniform Bayesian bba. When combining Bayesian bba's $m_1(.)$ and $m_2(.)$, the vacuous

prior and uniform prior $m_0(.)$ have therefore no impact on the result. Indeed, they contain no information that may help to prefer one particular state x_i with respect to the other ones, even if the level of conflict is different in both cases. So, the level of conflict doesn't matter at all in such Bayesian case. As already stated, what really matters is only the distribution of relative agreement factors. Only in such very particular cases (i.e. Bayesian bba's, and vacuous or Bayesian uniform priors), Dempster's rule is fully consistent with Bayes fusion rule.

B. Case 2: Non uniform Bayesian prior

Let us consider Dempster's fusion of Bayesian bba's with a Bayesian non uniform prior $m_0(.)$. In such case it is easy to check from the general structures of Bayes fusion rule and Dempster's fusion rule that these two rules are incompatible. Indeed, in Bayes rule one divides each posterior source $m_i(x_j)$ by $\sqrt[s]{m_0(x_j)}$, $i = 1, 2, \ldots s$, whereas the prior source $m_0(.)$ is combined in a pure conjunctive manner by Dempster's rule with the bba's $m_i(.)$, $i = 1, 2, \ldots s$, as if $m_0(.)$ was a simple additional source. This difference of processing prior information between the two approaches explains clearly the incompatibility of Dempster's rule with Bayes rule when Bayesian prior bba is not uniform. This incompatibility is illustrated in the next simple example.

Example 3: Let us consider the same frame $\Theta(X)$, and same bba's $m_1(.)$ and $m_2(.)$ as in the Example 3. Suppose that the prior information is Bayesian and non uniform as follows: $m_0(x_1) = P(X = x_1) = 0.6$, $m_0(x_2) = P(X = x_2) = 0.3$ and $m_0(x_3) = P(X = x_3) = 0.1$. Bayes rule (10) yields:

$$\begin{cases} P(x_1|Z_1 \cap Z_2) &= \frac{A_2(x_1)}{GA_2} = \frac{0.2 \cdot 0.5/0.6}{2.2667} = \frac{0.1667}{2.2667} \approx 0.0735 \\ P(x_2|Z_1 \cap Z_2) &= \frac{A_2(x_2)}{GA_2} = \frac{0.3 \cdot 0.1/0.3}{2.2667} = \frac{0.1000}{2.2667} \approx 0.0441 \\ P(x_3|Z_1 \cap Z_2) &= \frac{A_2(x_3)}{GA_2} = \frac{0.5 \cdot 0.4/0.1}{2.2667} = \frac{2.0000}{2.2667} \approx 0.8824 \end{cases}$$

Dempster's rule yields $m_{DS}(x_i) \neq P(x_i|Z_1 \cap Z_2)$ because:

$$\begin{cases} m_{DS}(x_1) &= \frac{1}{1-0.9110} \cdot 0.2 \cdot 0.5 \cdot 0.6 = \frac{0.060}{0.089} \approx 0.6742 \\ m_{DS}(x_2) &= \frac{1}{1-0.9110} \cdot 0.3 \cdot 0.1 \cdot 0.3 = \frac{0.009}{0.089} \approx 0.1011 \\ m_{DS}(x_3) &= \frac{1}{1-0.9110} \cdot 0.5 \cdot 0.4 \cdot 0.1 = \frac{0.020}{0.089} \approx 0.2247 \end{cases}$$

Therefore, one has in general:

$$DS(m_1(.), \ldots, m_s(.); m_0(.)) \neq Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X))$$

V. Conclusions

In this paper[1] we have analyzed in details the expression and the properties of Bayes rule of combination based on statistical conditional independence assumption, as well as the emblematic Dempster's rule of combination of belief functions introduced by Shafer in his Mathematical Theory of evidence. We have clearly explained from a theoretical standpoint, and also on simple examples, why Dempster's rule is not a generalization of Bayes rule in general. The incompatibility of Dempster's rule with Bayes rule is due to its impossibility to deal with non uniform Bayesian priors in the same manner as Bayes rule does. Dempster's rule turns to be compatible with Bayes rule only in two very particular cases: 1) if all the Bayesian bba's to combine (including the prior) focus on same state (i.e. there is a perfect conjunctive consensus between the sources), or 2) if all the bba's to combine (excluding the prior) are Bayesian, and if the prior bba cannot help to discriminate a particular state of the frame of discernment (i.e. the prior bba is either vacuous, or Bayesian and uniform). Except in these two very particular cases, Dempster's rule is totally incompatible with Bayes rule. Therefore, Dempster's rule cannot be claimed to be a generalization of Bayes fusion rule, even when the bba's to combine are Bayesian.

[1] An extended version of this paper will be presented at Fusion 2013 conference [20].

Acknowledgment

This study was partly supported by the project AComIn, grant 316087, funded by the FP7 Capacity Programme and co-supported by Grant for State Key Program for Basic Research of China (973) (No. 2013CB329405), National Natural Science Foundation of China (No.61104214, No. 61203222)

References

[1] L.A. Zadeh, On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, CA, U.S.A., 1979.

[2] A. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist., Vol. 38, pp. 325–339, 1967.

[3] A. Dempster, A generalization of bayesian inference, J. R. Stat. Soc. B 30, pp. 205–247, 1968.

[4] G. Shafer, A Mathematical theory of evidence, Princeton University Press, Princeton, NJ, U.S.A., 1976.

[5] L.A. Zadeh, Book review: A mathematical theory of evidence, The Al Magazine, Vol. 5, No. 3, pp. 81-83, 1984.

[6] L.A. Zadeh, A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The Al Magazine, Vol. 7, No. 2, pp. 85–90, 1986.

[7] J. Lemmer, Confidence factors, empiricism and the Dempster-Shafer theory of evidence, Proc. of UAI '85, pp. 160–176, Los Angeles, CA, U.S.A., July 10–12, 1985.

[8] G. Shafer, Perspectives on the theory and practice of belief functions, IJAR, Vol. 4, No. 5–6, pp. 323–362, 1990.

[9] J. Pearl, Reasoning with belief functions: an analysis of compatibility, IJAR, Vol. 4, No. 5–6, pp. 363–390, 1990.

[10] J. Pearl, Rejoinder of comments on "Reasoning with belief functions: An analysis of compatibility", IJAR, Vol. 6, No. 3, pp. 425–443, 1992.

[11] G.M. Provan, The validity of Dempster-Shafer belief functions, IJAR, Vol. 6., No. 3, pp. 389–399, May 1992.

[12] P. Wang, A defect in Dempster-Shafer theory, Proc. of UAI '94, pp. 560–566, Seattle, WA, U.S.A., July 29–31, 1994.

[13] A. Gelman, The boxer, the wrestler, and the coin flip: a paradox of robust bayesian inference and belief functions, American Statistician, Vol. 60, No. 2, pp. 146–150, 2006.

[14] F. Smarandache, J. Dezert (Editors), Applications and advances of DSmT for information fusion, Vol. 3, ARP, U.S.A., 2009. http://fs.gallup.unm.edu/DSmT.htm

[15] J. Dezert, P. Wang, A. Tchamova, On the validity of Dempster-Shafer theory, Proc. of Fusion 2012 Int. Conf., Singapore, July 9–12, 2012.

[16] A. Tchamova, J. Dezert, On the Behavior of Dempster's rule of combination and the foundations of Dempster-Shafer theory, (best paper award), Proc. of 6th IEEE Int. Conf. on Intelligent Systems IS '12, Sofia, Bulgaria, Sept. 6–8, 2012.

[17] R.P. Mahler, Statistical multisource-multitarget information fusion, Chapter 4, Artech House, 2007.

[18] X.R. Li, Probability, random signals and statistics, CRC Press, 1999.

[19] G. Shafer, Non-additive probabilities in the work of Bernoulli and Lambert, in Archive for History of Exact Sciences, C. Truesdell (Ed.), Springer-Verlag, Berlin, Vol. 19, No. 4, pp. 309–370, 1978.

[20] J. Dezert, A. Tchamova, D. Han, J. Tacnet, Why Dempsters fusion rule is not a generalization of Bayes fusion rule, Proc. of Fusion 2013 Int. Conf., Istanbul, 2013 (accepted for publication).

Why Dempster's Fusion Rule is not a Generalization of Bayes Fusion Rule

Jean Dezert
Albena Tchamova
Deqiang Han
Jean-Marc Tacnet

Originally published as Dezert J., Tchamova A., Han D., Tacnet J.-M., *Why Dempster's fusion rule is not a generalization of Bayes fusion rule*, Proc. Of Fusion 2013 Int. Conference on Information Fusion, Istanbul, Turkey, July 9-12, 2013, and reprinted with permission.

***Abstract*—In this paper, we analyze Bayes fusion rule in details from a fusion standpoint, as well as the emblematic Dempster's rule of combination introduced by Shafer in his Mathematical Theory of evidence based on belief functions. We propose a new interesting formulation of Bayes rule and point out some of its properties. A deep analysis of the compatibility of Dempster's fusion rule with Bayes fusion rule is done. We show that Dempster's rule is compatible with Bayes fusion rule only in the very particular case where the basic belief assignments (bba's) to combine are Bayesian, and when the prior information is modeled either by a uniform probability measure, or by a vacuous bba. We show clearly that Dempster's rule becomes incompatible with Bayes rule in the more general case where the prior is truly informative (not uniform, nor vacuous). Consequently, this paper proves that Dempster's rule is not a generalization of Bayes fusion rule.**

***Keywords*—Information fusion, Probability theory, Bayes fusion rule, Dempster's fusion rule.**

I. Introduction

In 1979, Lotfi Zadeh questioned in [1] the validity of the Dempster's rule of combination [2], [3] proposed by Shafer in Dempster-Shafer Theory (DST) of evidence [4]. Since more than 30 years many strong debates [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15] on the validity of foundations of DST and Dempster's rule have bloomed. The purpose of this paper is not to discuss the validity of Dempster's rule, nor the foundations of DST which have been already addressed in previous papers [16], [17], [18]. In this paper, we just focus on the deep analysis of the real incompatibility of Dempster's rule with Bayes fusion rule. Our analysis supports Mahler's one briefly presented in [19].

This paper is organized as follows. In section II, we recall basics of conditional probabilities and Bayes fusion rule with its main properties. In section III, we recall the basics of belief functions and Dempster's rule. In section IV, we analyze in details the incompatibility of Dempster's rule with Bayes rule in general and its partial compatibility for the very particular case when prior information is modeled by a Bayesian uniform basic belief assignment (bba). Section V concludes this paper.

II. Conditional probabilities and Bayes fusion

In this section, we recall the definition of conditional probability [20], [21] and present the principle and the properties of Bayes fusion rule. We present the structure of this rule derived from the classical definition of the conditional probability in a new uncommon interesting form that will help us to analyze its partial similarity with Dempster's rule proposed by Shafer in his mathematical theory of evidence [4]. We will show clearly why Dempster's rule fails to be compatible with Bayes rule in general.

A. Conditional probabilities

Let us consider two random events X and Z. The conditional probability mass functions (pmfs) $P(X|Z)$ and $P(Z|X)$ are defined[1] (assuming $P(X) > 0$ and $P(Z) > 0$) by [20]:

$$P(X|Z) \triangleq \frac{P(X \cap Z)}{P(Z)} \quad \text{and} \quad P(Z|X) \triangleq \frac{P(X \cap Z)}{P(X)} \tag{1}$$

From Eq. (1), one gets $P(X \cap Z) = P(X|Z)P(Z) = P(Z|X)P(X)$, which yields to Bayes Theorem:

$$P(X|Z) = \frac{P(Z|X)P(X)}{P(Z)} \text{ and } P(Z|X) = \frac{P(X|Z)P(Z)}{P(X)} \tag{2}$$

where $P(X)$ is called the *a priori probability* of X, and $P(Z|X)$ is called the *likelihood of* X. The denominator $P(Z)$ plays the role of a normalization constant warranting that $\sum_{i=1}^{N} P(X = x_i|Z) = 1$. In fact $P(Z)$ can be rewritten as

$$P(Z) = \sum_{i=1}^{N} P(Z|X = x_i)P(X = x_i) \tag{3}$$

The set of the N possible exclusive and exhaustive outcomes of X is denoted $\Theta(X) \triangleq \{x_i, i = 1, 2, \ldots, N\}$.

B. Bayes parallel fusion rule

In fusion applications, we are often interested in computing the probability of an event X given two events Z_1 and Z_2 that have occurred. More precisely, one wants to compute $P(X|Z_1 \cap Z_2)$ knowing $P(X|Z_1)$ and $P(X|Z_2)$, where X can take N distinct exhaustive and exclusive states x_i, $i = 1, 2, \ldots, N$. Such type of problem is traditionally called a *fusion problem*. The computation of $P(X|Z_1 \cap Z_2)$ from

[1] For convenience and simplicity, we use the notation $P(X|Z)$ instead of $P(X = x|Z = z)$, and $P(Z|X)$ instead of $P(Z = z|X = x)$ where x and z would represent precisely particular outcomes of the random variables X and Z.

$P(X|Z_1)$ and $P(X|Z_2)$ cannot be done in general without the knowledge of the probabilities $P(X)$ and $P(X|Z_1 \cup Z_2)$ which are rarely given. However, $P(X|Z_1 \cap Z_2)$ becomes easily computable by assuming the following conditional statistical independence condition expressed mathematically by:

$$(A1): \qquad P(Z_1 \cap Z_2|X) = P(Z_1|X)P(Z_2|X) \tag{4}$$

With such conditional independence condition (A1), then from Eq. (1) and Bayes Theorem one gets:

$$\begin{aligned} P(X|Z_1 \cap Z_2) &= \frac{P(Z_1 \cap Z_2 \cap X)}{P(Z_1 \cap Z_2)} = \frac{P(Z_1 \cap Z_2|X)P(X)}{P(Z_1 \cap Z_2)} \\ &= \frac{P(Z_1|X)P(Z_2|X)P(X)}{\sum_{i=1}^{N} P(Z_1|X=x_i)P(Z_2|X=x_i)P(X=x_i)} \end{aligned}$$

Using again Eq. (2), we have:

$$P(Z_1|X) = \frac{P(X|Z_1)P(Z_1)}{P(X)} \text{ and } P(Z_2|X) = \frac{P(X|Z_2)P(Z_2)}{P(X)}$$

and the previous formula of conditional probability $P(X|Z_1 \cap Z_2)$ can be rewritten as:

$$P(X|Z_1 \cap Z_2) = \frac{\frac{P(X|Z_1)P(X|Z_2)}{P(X)}}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)}} \tag{5}$$

The rule of combination given by Eq. (5) is known as *Bayes parallel (or product) rule* and dates back to Bernoulli [22]. In the classification framework, this formula is also called the *Naive Bayesian Classifier* because it uses the assumption (A1) which is often considered as very unrealistic and too simplistic, and that is why it is called a *naive* assumption. The Eq. (5) can be rewritten as:

$$P(X|Z_1 \cap Z_2) = \frac{1}{K(X, Z_1, Z_2)} \cdot P(X|Z_1) \cdot P(X|Z_2) \tag{6}$$

where the coefficient $K(X, Z_1, Z_2)$ is defined by:

$$K(X, Z_1, Z_2) \triangleq P(X) \cdot \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_2)}{P(X=x_i)} \tag{7}$$

C. Symmetrization of Bayes fusion rule

The expression of Bayes fusion rule given by Eq. (5) can also be symmetrized in the following form that, quite surprisingly, rarely appears in the literature:

$$P(X|Z_1 \cap Z_2) = \frac{\frac{P(X|Z_1)}{\sqrt{P(X)}} \cdot \frac{P(X|Z_2)}{\sqrt{P(X)}}}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}}} \tag{8}$$

or in an equivalent manner:

$$P(X|Z_1 \cap Z_2) = \frac{1}{K'(Z_1, Z_2)} \cdot \frac{P(X|Z_1)}{\sqrt{P(X)}} \cdot \frac{P(X|Z_2)}{\sqrt{P(X)}} \tag{9}$$

where the normalization constant $K'(Z_1, Z_2)$ is given by:

$$K'(Z_1, Z_2) \triangleq \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}} \tag{10}$$

We call the quantity $A_2(X = x_i) \triangleq \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}}$ entering in Eq. (10) the *Agreement Factor* on $X = x_i$ of order 2, because only two posterior pmfs are used in the derivation. $A_2(X = x_i)$ corresponds to the posterior conjunctive consensus on the event $X = x_i$ taking into account the prior pmf of X. The denominator of Eq. (8) measures the level of the Global Agreement (GA) of the conjunctive consensus taking into account the prior pmf of X. It is denoted[2] GA_2.

$$\begin{aligned} GA_2 &\triangleq \sum_{i_1,i_2=1|i_1=i_2}^{N} \frac{P(X=x_{i_1}|Z_1)}{\sqrt{P(X=x_{i_1})}} \cdot \frac{P(X=x_{i_2}|Z_2)}{\sqrt{P(X=x_{i_2})}} \\ &= \sum_{i=1}^{N} \frac{P(X=x_i|Z_1)}{\sqrt{P(X=x_i)}} \cdot \frac{P(X=x_i|Z_2)}{\sqrt{P(X=x_i)}} = K'(Z_1, Z_2) \end{aligned} \tag{11}$$

In fact, with assumption (A1), the probability $P(X|Z_1 \cap Z_2)$ given in Eq. (9) is nothing but the simple ratio of the agreement factor $A_2(X)$ (conjunctive consensus) on X over the global agreement $GA_2 = \sum_{i=1}^{N} A_2(X = x_i)$, that is:

$$P(X|Z_1 \cap Z_2) = \frac{A_2(X)}{GA_2} \tag{12}$$

The quantity GC_2 given in Eq. (13) measures the global conflict (i.e. the total conjunctive disagreement) taking into account the prior pmf of X.

$$GC_2 \triangleq \sum_{i_1,i_2=1|i_1 \neq i_2}^{N} \frac{P(X=x_{i_1}|Z_1)}{\sqrt{P(X=x_{i_1})}} \cdot \frac{P(X=x_{i_2}|Z_2)}{\sqrt{P(X=x_{i_2})}} \tag{13}$$

• **Generalization to** $P(X|Z_1 \cap Z_2 \cap \ldots \cap Z_s)$

It can be proved that, when assuming conditional independence conditions, Bayes parallel combination rule can be generalized for combining $s > 2$ posterior pmfs as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{1}{K(X, Z_1, \ldots, Z_s)} \cdot \prod_{k=1}^{s} P(X|Z_k) \tag{14}$$

where the coefficient $K(X, Z_1, \ldots, Z_s)$ is defined by:

$$K(X, Z_1, \ldots, Z_s) \triangleq P(X) \sum_{i=1}^{N} \frac{(\prod_{k=1}^{s} P(X=x_i|Z_k))}{P(X=x_i)} \tag{15}$$

The symmetrized form of Eq. (14) is:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{1}{K'(Z_1, \ldots, Z_s)} \cdot \prod_{k=1}^{s} \frac{P(X|Z_k)}{\sqrt[s]{P(X)}} \tag{16}$$

with the normalization constant $K'(Z_1, \ldots, Z_s)$ given by:

$$K'(Z_1, \ldots, Z_s) \triangleq \sum_{i=1}^{N} \prod_{k=1}^{s} \frac{P(X=x_i|Z_k)}{\sqrt[s]{P(X=x_i)}} \tag{17}$$

[2]The index 2 is introduced explicitly in the notations because we consider only the fusion of two posterior pmfs.

The generalization of $A_2(X)$, GA_2, and GC_2 provides the agreement $A_s(X)$ of order s, the global agreement GA_s and the global conflict GC_s for s sources as follows:

$$A_s(X=x_i) \triangleq \prod_{k=1}^{s} \frac{P(X=x_i|Z_k)}{\sqrt[s]{P(X=x_i)}}$$

$$GA_s \triangleq \sum_{i_1,\ldots,i_s=1|i_1=\ldots=i_s}^{N} \frac{P(X=x_{i_1}|Z_1)}{\sqrt[s]{P(X=x_{i_1})}} \cdots \frac{P(X=x_{i_s}|Z_s)}{\sqrt[s]{P(X=x_{i_s})}}$$

$$GC_s \triangleq \sum_{i_1,\ldots,i_s=1}^{N} \frac{P(X=x_{i_1}|Z_1)}{\sqrt[s]{P(X=x_{i_1})}} \cdots \frac{P(X=x_{i_s}|Z_s)}{\sqrt[s]{P(X=x_{i_s})}} - GA_s$$

• Symbolic representation of Bayes fusion rule

The (symmetrized form of) Bayes fusion rule of two posterior probability measures $P(X|Z_1)$ and $P(X|Z_2)$, given in Eq. (9), requires an extra knowledge of the prior probability of X. For convenience, we denote symbolically this fusion rule as:

$$P(X|Z_1 \cap Z_2) = Bayes(P(X|Z_1), P(X|Z_2); P(X)) \quad (18)$$

Similarly, the (symmetrized) Bayes fusion rule of $s \geq 2$ probability measures $P(X|Z_k)$, $k = 1, 2, \ldots, s$ given by Eq. (16), which requires also the knowledge of $P(X)$, will be denoted as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X))$$

• Particular case: Uniform a priori pmf

If the random variable X is assumed as a priori uniformly distributed over the space of its N possible outcomes, then the probability of X is equal to $P(X = x_i) = 1/N$ for $i = 1, 2, \ldots, N$. In such particular case, all the prior probabilities values $\sqrt{P(X=x_i)} = \sqrt{1/N}$ and $\sqrt[s]{P(X=x_i)} = \sqrt[s]{1/N}$ can be simplified in Bayes fusion formulas Eq. (9) and Eq. (10). Therefore, Bayes fusion formula (9) reduces to:

$$P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1)P(X|Z_2)}{\sum_{i=1}^{N} P(X=x_i|Z_1)P(X=x_i|Z_2)} \quad (19)$$

By convention, Eq. (19) is denoted symbolically as:

$$P(X|Z_1 \cap Z_2) = Bayes(P(X|Z_1), P(X|Z_2)) \quad (20)$$

Similarly, $Bayes(P(X|Z_1), \ldots, P(X|Z_s))$ rule defined with an uniform a priori pmf of X will be given by:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{\sum_{i=1}^{N} \prod_{k=1}^{s} P(X=x_i|Z_k)} \quad (21)$$

When $P(X)$ is uniform and from Eq. (19), one can redefine the global agreement and the global conflict as:

$$GA_2^{unif} \triangleq \sum_{i,j=1|i=j}^{N} P(X=x_i|Z_1)P(X=x_j|Z_2) \quad (22)$$

$$GC_2^{unif} \triangleq \sum_{i,j=1|i\neq j}^{N} P(X=x_i|Z_1)P(X=x_j|Z_2) \quad (23)$$

Because $\sum_{i=1}^{N} P(X = x_i|Z_1) = 1$ and $\sum_{j=1}^{N} P(X = x_j|Z_2) = 1$, then

$$\begin{aligned} 1 &= (\sum_{i=1}^{N} P(X=x_i|Z_1))(\sum_{j=1}^{N} P(X=x_j|Z_2)) \\ &= \sum_{i,j=1}^{N} P(X=x_i|Z_1)P(X=x_j|Z_2) \\ &= \sum_{i,j=1|i=j}^{N} P(X=x_i|Z_1)P(X=x_j|Z_2) \\ &\quad + \sum_{i,j=1|i\neq j}^{N} P(X=x_i|Z_1)P(X=x_j|Z_2) \end{aligned}$$

Therefore, one has always $GA_2^{unif} + GC_2^{unif} = 1$ when $P(X)$ is uniform, and Eq. (19) can be expressed as:

$$P(X|Z_1 \cap Z_2) = \frac{P(X|Z_1)P(X|Z_2)}{GA_2^{unif}} = \frac{P(X|Z_1)P(X|Z_2)}{1 - GC_2^{unif}} \quad (24)$$

By a direct extension, one will have:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\prod_{k=1}^{s} P(X|Z_k)}{GA_s^{unif}} = \frac{\prod_{k=1}^{s} P(X|Z_k)}{1 - GC_s^{unif}} \quad (25)$$

$$GA_s^{unif} = \sum_{i_1,\ldots,i_s=1|i_1=\ldots=i_s}^{N} P(X=x_{i_1}|Z_1)\ldots P(X=x_{i_s}|Z_s)$$

$$GC_s^{unif} = 1 - GA_s^{unif}$$

Remark 1: The normalization coefficient corresponding to the global conjunctive agreement GA_s^{unif} can also be expressed using belief function notations [4] as:

$$GA_s^{unif} = \sum_{\substack{x_{i_1},\ldots,x_{i_s}\in\Theta(X) \\ x_{i_1}\cap\ldots\cap x_{i_s}\neq\emptyset}} P(X=x_{i_1}|Z_1)\ldots P(X=x_{i_s}|Z_s)$$

and the global disagreement, or total conflict level, is given by:

$$GC_s^{unif} = \sum_{\substack{x_{i_1},\ldots,x_{i_s}\in\Theta(X) \\ x_{i_1}\cap\ldots\cap x_{i_s}=\emptyset}} P(X=x_{i_1}|Z_1)\ldots P(X=x_{i_s}|Z_s)$$

D. Properties of Bayes fusion rule

In this subsection, we analyze Bayes fusion rule (assuming condition (A1) holds) from a pure algebraic standpoint. In fusion jargon, the quantities to combine come from sources of information which provide inputs that feed the fusion rule. In the probabilistic framework, a source s to combine corresponds to the posterior pmf $P(X|Z_s)$. In this subsection, we establish five interesting properties of Bayes rule. Contrary to Dempster's rule, we prove that Bayes rule is not associative in general.

• (P1) : The pmf $P(X)$ is a neutral element of Bayes fusion rule when combining only two sources.

Proof: A source is called a neutral element of a fusion rule if and only if it has no influence on the fusion result. $P(X)$ is a neutral element of Bayes rule if and only if

$Bayes(P(X|Z_1), P(X); P(X)) = P(X|Z_1)$. It can be easily verified that this equality holds by replacing $P(X|Z_2)$ by $P(X)$ and $P(X = x_i|Z_2)$ by $P(X = x_i)$ (as if the conditioning term Z_2 vanishes) in Eq. (5). One can also verify that $Bayes(P(X), P(X|Z_2); P(X)) = P(X|Z_2)$, which completes the proof.

Remark 2: When considering Bayes fusion of more than two sources, $P(X)$ doesn't play the role of a neutral element in general, except if $P(X)$ is uniform. For example, let us consider 3 pmfs $P(X|Z_1)$, $P(X|Z_2)$ and $P(X|Z_3)$ to combine with formula (14) with $P(X)$ not uniform. When Z_3 vanishes so that $P(X|Z_3) = P(X)$, we can easily check that:

$$Bayes(P(X|Z_1), P(X|Z_2), P(X); P(X)) \neq Bayes(P(X|Z_1), P(X|Z_2); P(X)) \quad (26)$$

• (P2) : Bayes fusion rule is in general not idempotent.

Proof: A fusion rule is idempotent if the combination of all same inputs is equal to the inputs. To prove that Bayes rule is not idempotent it suffices to prove that in general:

$$Bayes(P(X|Z_1), P(X|Z_1); P(X)) \neq P(X|Z_1)$$

From Bayes rule (5), when $P(X|Z_2) = P(X|Z_1)$ we clearly get in general

$$\frac{1}{P(X)} \frac{P(X|Z_1)P(X|Z_1)}{\sum_{i=1}^{N} \frac{P(X=x_i|Z_1)P(X=x_i|Z_1)}{P(X=x_i)}} \neq P(X|Z_1) \quad (27)$$

but when Z_1 and Z_2 vanish, because in such case Eq. (27) reduces to $P(X)$ on its left and right sides.

Remark 3: In the particular (two sources) degenerate case where Z_1 and Z_2 vanish, one has always: $Bayes(P(X), P(X); P(X)) = P(X)$. However, in the more general degenerate case (when considering more than 2 sources), one will have in general: $Bayes(P(X), P(X), \ldots, P(X); P(X)) \neq P(X)$, but when $P(X)$ is uniform, or when $P(X)$ is a "deterministic" probability measure such that $P(X = x_i) = 1$ for a given $x_i \in \Theta(X)$ and $P(X = x_j) = 0$ for all $x_j \neq x_i$.

• (P3) : Bayes fusion rule is in general not associative.

Proof: A fusion rule f is called associative if and only if it satisfies the associative law: $f(f(x, y), z) = f(x, f(y, z)) = f(y, f(x, z)) = f(x, y, z)$ for all possible inputs x, y and z. Let us prove that Bayes rule is not associative from a very simple example.

Example 1: Let us consider the simplest set of outcomes $\{x_1, x_2\}$ for X, with prior pmf:

$$P(X = x_1) = 0.2 \text{ and } P(X = x_2) = 0.8$$

and let us consider the three given sets of posterior pmfs:

$$\begin{cases} P(X = x_1|Z_1) = 0.1 \text{ and } P(X = x_2|Z_1) = 0.9 \\ P(X = x_1|Z_2) = 0.5 \text{ and } P(X = x_2|Z_2) = 0.5 \\ P(X = x_1|Z_3) = 0.6 \text{ and } P(X = x_2|Z_3) = 0.4 \end{cases}$$

Bayes fusion $Bayes(P(X|Z_1),)P(X|Z_2), P(X|Z_3); P(X))$ of the three sources altogether according to Eq. (16) provides:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_2 \cap Z_3) = \frac{1}{K_{123}} \frac{0.1}{\sqrt[3]{0.2}} \frac{0.5}{\sqrt[3]{0.2}} \frac{0.6}{\sqrt[3]{0.2}} = 0.40 \\ P(X = x_2|Z_1 \cap Z_2 \cap Z_3) = \frac{1}{K_{123}} \frac{0.9}{\sqrt[3]{0.8}} \frac{0.5}{\sqrt[3]{0.8}} \frac{0.4}{\sqrt[3]{0.8}} = 0.60 \end{cases}$$

where the normalization constant K_{123} is given by:

$$K_{123} = \frac{0.1}{\sqrt[3]{0.2}} \frac{0.5}{\sqrt[3]{0.2}} \frac{0.6}{\sqrt[3]{0.2}} + \frac{0.9}{\sqrt[3]{0.8}} \frac{0.5}{\sqrt[3]{0.8}} \frac{0.4}{\sqrt[3]{0.8}} = 0.3750$$

Let us compute the fusion of $P(X|Z_1)$ with $P(X|Z_2)$ using $Bayes(P(X|Z_1), P(X|Z_2); P(X))$. One has:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_2) = \frac{1}{K_{12}} \frac{0.1}{\sqrt{0.2}} \frac{0.5}{\sqrt{0.2}} \approx 0.3077 \\ P(X = x_2|Z_1 \cap Z_2) = \frac{1}{K_{12}} \frac{0.9}{\sqrt{0.8}} \frac{0.5}{\sqrt{0.8}} \approx 0.6923 \end{cases}$$

where the normalization constant K_{12} is given by:

$$K_{12} = \frac{0.1}{\sqrt{0.2}} \frac{0.5}{\sqrt{0.2}} + \frac{0.9}{\sqrt{0.8}} \frac{0.5}{\sqrt{0.8}} = 0.8125$$

Let us compute the fusion of $P(X|Z_2)$ with $P(X|Z_3)$ using $Bayes(P(X|Z_2), P(X|Z_3); P(X))$. One has

$$\begin{cases} P(X = x_1|Z_2 \cap Z_3) = \frac{1}{K_{23}} \frac{0.5}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} \approx 0.8571 \\ P(X = x_2|Z_2 \cap Z_3) = \frac{1}{K_{23}} \frac{0.5}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} \approx 0.1429 \end{cases}$$

where the normalization constant K_{23} is given by:

$$K_{23} = \frac{0.5}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} + \frac{0.5}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} = 1.75$$

Let us compute the fusion of $P(X|Z_1)$ with $P(X|Z_3)$ using $Bayes(P(X|Z_1), P(X|Z_3); P(X))$. One has:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_3) = \frac{1}{K_{13}} \frac{0.1}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} = 0.4 \\ P(X = x_2|Z_1 \cap Z_3) = \frac{1}{K_{13}} \frac{0.9}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} = 0.6 \end{cases}$$

where the normalization constant K_{13} is given by:

$$K_{13} = \frac{0.1}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} + \frac{0.9}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} = 0.75$$

Let us compute the fusion of $P(X|Z_1 \cap Z_2)$ with $P(X|Z_3)$ using $Bayes(P(X|Z_1 \cap Z_2), P(X|Z_3); P(X))$. One has

$$\begin{cases} P(X = x_1|(Z_1 \cap Z_2) \cap Z_3) = \frac{1}{K_{(12)3}} \frac{0.3077}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} \approx 0.7273 \\ P(X = x_2|(Z_1 \cap Z_2) \cap Z_3) = \frac{1}{K_{(12)3}} \frac{0.6923}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} \approx 0.2727 \end{cases}$$

where the normalization constant $K_{(12)3}$ is given by

$$K_{(12)3} = \frac{0.3077}{\sqrt{0.2}} \frac{0.6}{\sqrt{0.2}} + \frac{0.6923}{\sqrt{0.8}} \frac{0.4}{\sqrt{0.8}} \approx 1.26925$$

Let us compute the fusion of $P(X|Z_1)$ with $P(X|Z_2 \cap Z_3)$ using $Bayes(P(X|Z_1), P(X|Z_2 \cap Z_3); P(X))$. One has

$$\begin{cases} P(X = x_1|Z_1 \cap (Z_2 \cap Z_3)) = \frac{1}{K_{1(23)}} \frac{0.1}{\sqrt{0.2}} \frac{0.8571}{\sqrt{0.2}} \approx 0.7273 \\ P(X = x_2|Z_1 \cap (Z_2 \cap Z_3)) = \frac{1}{K_{1(23)}} \frac{0.9}{\sqrt{0.8}} \frac{0.1429}{\sqrt{0.8}} \approx 0.2727 \end{cases}$$

where the normalization constant $K_{1(23)}$ is given by

$$K_{1(23)} = \frac{0.1}{\sqrt{0.2}} \frac{0.8571}{\sqrt{0.2}} + \frac{0.9}{\sqrt{0.8}} \frac{0.1429}{\sqrt{0.8}} \approx 0.58931$$

Let us compute the fusion of $P(X|Z_1 \cap Z_3)$ with $P(X|Z_2)$ using $Bayes(P(X|Z_1 \cap Z_3), P(X|Z_2); P(X))$. One has

$$\begin{cases} P(X = x_1|(Z_1 \cap Z_3) \cap Z_2) = \frac{1}{K_{(13)2}} \frac{0.4}{\sqrt{0.2}} \frac{0.5}{\sqrt{0.2}} \approx 0.7273 \\ P(X = x_2|(Z_1 \cap Z_3) \cap Z_2) = \frac{1}{K_{(13)2}} \frac{0.6}{\sqrt{0.8}} \frac{0.5}{\sqrt{0.8}} \approx 0.2727 \end{cases}$$

where the normalization constant $K_{(13)2}$ is given by

$$K_{(13)2} = \frac{0.4}{\sqrt{0.2}}\frac{0.5}{\sqrt{0.2}} + \frac{0.6}{\sqrt{0.8}}\frac{0.5}{\sqrt{0.8}} = 1.375$$

Therefore, one sees that even if in our example one has $f(x, f(y, z)) = f(f(x, y), z) = f(y, f(x, z))$ because $P(X|(Z_1 \cap Z_2) \cap Z_3) = P(X|Z_1 \cap (Z_2 \cap Z_3)) = P(X|Z_2 \cap (Z_1 \cap Z_3))$, Bayes fusion rule is not associative since:

$$\begin{cases} P(X|(Z_1 \cap Z_2) \cap Z_3) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \\ P(X|Z_1 \cap (Z_2 \cap Z_3)) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \\ P(X|Z_2 \cap (Z_1 \cap Z_3)) \neq P(X|Z_1 \cap Z_2 \cap Z_3) \end{cases}$$

• (P4) : Bayes fusion rule is associative if and only if $P(X)$ is uniform.

Proof: If $P(X)$ is uniform, Bayes fusion rule is given by Eq. (21) which can be rewritten as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{P(X|Z_s)\prod_{k=1}^{s-1} P(X|Z_k)}{\sum_{i=1}^{N} P(X = x_i|Z_s)\prod_{k=1}^{s-1} P(X = x_i|Z_k)}$$

By introducing the term $1/\sum_{i=1}^{N}\prod_{k=1}^{s-1} P(X = x_i|Z_k)$ in numerator and denominator of the previous formula, it comes:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{\frac{\prod_{k=1}^{s-1} P(X|Z_k)}{\sum_{i=1}^{N}\prod_{k=1}^{s-1} P(X=x_i|Z_k)} P(X|Z_s)}{\sum_{i=1}^{N} \frac{\prod_{k=1}^{s-1} P(X=x_i|Z_k)}{\sum_{i=1}^{N}\prod_{k=1}^{s-1} P(X=x_i|Z_k)} P(X = x_i|Z_s)}$$

which can be simply rewritten as:

$$P(X|Z_1 \cap \ldots \cap Z_s) = \frac{P(X|Z_1 \cap \ldots \cap Z_{s-1})P(X|Z_s)}{\sum_{i=1}^{N} P(X = x_i|Z_1 \cap \ldots \cap Z_{s-1})P(X = x_i|Z_s)}$$

Therefore when $P(X)$ is uniform, one has:

$$\begin{aligned} &Bayes(P(X|Z_1), \ldots, P(X|Z_s)) = \\ &\quad Bayes(Bayes(P(X|Z_1), \ldots, P(X|Z_{s-1})), P(X|Z_s)) \end{aligned}$$

The previous relation was based on the decomposition of $\prod_{k=1}^{s} P(X|Z_k)$ as $P(X|Z_s)\prod_{k=1}^{s-1} P(X|Z_k)$. This choice of decomposition was arbitrary and chosen only for convenience. In fact $\prod_{k=1}^{s} P(X|Z_k)$ can be decomposed in s different manners, as $P(X|Z_j)\prod_{k=1|k\neq j}^{s} P(X|Z_k)$, $j = 1, 2, \ldots s$ and the similar analysis can be done. In particular, when $s = 3$, we will have:

$$\begin{aligned} &Bayes(P(X|Z_1), P(X|Z_2), P(X|Z_3)) = \\ &\quad Bayes(Bayes(P(X|Z_1), P(X|Z_2)), P(X|Z_3)) \\ &\quad = Bayes(P(X|Z_1), Bayes(P(X|Z_2), P(X|Z_3))) \end{aligned}$$

which completes the proof.

• (P5) : The levels of global agreement and global conflict between the sources do not matter in Bayes fusion rule.

Proof: This property seems surprising at first glance, but, since the results of Bayes fusion is nothing but the ratio of the agreement on x_i ($i = 1, 2, \ldots, N$) over the global agreement factor, many distinct sources with different global agreements (and thus with different global conflicts) can yield same Bayes fusion result. Indeed, the ratio is kept unchanged when multiplying its numerator and denominator by same non null scalar value. Consequently, the absolute levels of global agreement between the sources (and therefore of global conflict also) do not matter in Bayes fusion result. What really matters is only the proportions of relative agreement factors.

Example 2: To illustrate this property, let us consider Bayes fusion rule applied to two distinct sets[3] of sources represented by $Bayes(P(X|Z_1), P(X|Z_2); P(X))$ and by $Bayes(P'(X|Z_1), P'(X|Z_2); P(X))$ with the following prior and posterior pmfs:

$$P(X = x_1) = 0.2 \text{ and } P(X = x_2) = 0.8$$

$$\begin{cases} P(X = x_1|Z_1) \approx 0.0607 \text{ and } P(X = x_2|Z_1) \approx 0.9393 \\ P(X = x_1|Z_2) \approx 0.6593 \text{ and } P(X = x_2|Z_2) \approx 0.3407 \end{cases}$$

$$\begin{cases} P'(X = x_1|Z_1) \approx 0.8360 \text{ and } P'(X = x_2|Z_1) \approx 0.1640 \\ P'(X = x_1|Z_2) \approx 0.0240 \text{ and } P'(X = x_2|Z_2) \approx 0.9760 \end{cases}$$

Applying Bayes fusion rule given by Eq. (5), one gets for $Bayes(P(X|Z_1), P(X|Z_2); P(X))$:

$$\begin{cases} P(X = x_1|Z_1 \cap Z_2) = \frac{0.2}{0.2+0.4} = 1/3 \\ P(X = x_2|Z_1 \cap Z_2) = \frac{0.4}{0.2+0.4} = 2/3 \end{cases} \quad (28)$$

Similarly, one gets for $Bayes(P'(X|Z_1), P'(X|Z_2); P(X))$

$$\begin{cases} P'(X = x_1|Z_1 \cap Z_2) = \frac{0.1}{0.1+0.2} = 1/3 \\ P'(X = x_2|Z_1 \cap Z_2) = \frac{0.2}{0.1+0.2} = 2/3 \end{cases} \quad (29)$$

Therefore, one sees that $Bayes(P(X|Z_1), P(X|Z_2); P(X)) = Bayes(P'(X|Z_1), P'(X|Z_2); P(X))$ even if the levels of global agreements (and global conflicts) are different. In this particular example, one has:

$$\begin{cases} (GA_2 = 0.60) \neq (GA'_2 = 0.30) \\ (GC_2 = 1.60) \neq (GC'_2 = 2.05) \end{cases} \quad (30)$$

In summary, different sets of sources to combine (with different levels of global agreement and global conflict) can provide exactly the same result once combined with Bayes fusion rule. Hence the different levels of global agreement and global conflict do not really matter in Bayes fusion rule. What really matters in Bayes fusion rule is only the distribution of all the relative agreement factors defined as $A_s(X = x_i)/GA_s$.

III. Belief functions and Dempster's rule

The Belief Functions (BF) have been introduced in 1976 by Glenn Shafer in his mathematical theory of evidence [4], also known as Dempster-Shafer Theory (DST) in order to reason under uncertainty and to model epistemic uncertainties. We will not present in details the foundations of DST, but only the basic mathematical definitions that are necessary for the scope of this paper. The emblematic fusion rule proposed by Shafer to combine sources of evidences characterized by their basic belief assignments (bba) is Dempster's rule that will be analyzed in details in the sequel. In the literature over the years, DST has been widely defended by its proponents in arguing that: 1) Probability measures are particular cases of Belief

[3]The values chosen for $P(X|Z_1)$, $P(X|Z_2)$, $P'(X|Z_1)$, $P'(X|Z_2)$ here have been approximated at the fourth digit. They can be precisely determined such that the expressions for $P(X|Z_1 \cap Z_2)$ and $P'(X|Z_1 \cap Z_2)$ as given in Eqs. (28) and (29) hold. For example, the exact value of $P(x_1|Z_2)$ is obtained by solving a polynomial equation of degree 2 having as a possible solution $P(x_1|Z_2) = \frac{1}{2}(0.72 + \sqrt{0.72^2 - 4 \times 0.04}) = 0.659332590941915 \approx 0.6593$, etc.

functions; and 2) Dempster's fusion rule is a generalization of Bayes fusion rule. Although the statement 1) is correct because Probability measures are indeed particular (additive) Belief functions (called as Bayesian belief functions), we will explain why the second statement about Dempster's rule is incorrect in general.

A. Belief functions

Let Θ be a frame of discernment of a problem under consideration. More precisely, the set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ consists of a list of N exhaustive and exclusive elements θ_i, $i = 1, 2, \ldots, N$. Each θ_i represents a possible state related to the problem we want to solve. The exhaustivity and exclusivity of elements of Θ is referred as Shafer's model of the frame Θ. A basic belief assignment (bba), also called a belief mass function, $m(.) : 2^\Theta \to [0, 1]$ is a mapping from the power set of Θ (i.e. the set of subsets of Θ), denoted 2^Θ, to $[0, 1]$, that verifies the following conditions [4]:

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{X \in 2^\Theta} m(X) = 1 \tag{31}$$

The quantity $m(X)$ represents the mass of belief exactly committed to X. An element $X \in 2^\Theta$ is called a focal element if and only if $m(X) > 0$. The set $\mathcal{F}(m) \triangleq \{X \in 2^\Theta | m(X) > 0\}$ of all focal elements of a bba $m(.)$ is called the core of the bba. A bba $m(.)$ is said Bayesian if its focal elements are singletons of 2^Θ. The vacuous bba characterizing the total ignorance denoted[4] $I_t = \theta_1 \cup \theta_2 \cup \ldots \cup \theta_N$ is defined by $m_v(.) : 2^\Theta \to [0; 1]$ such that $m_v(X) = 0$ if $X \neq \Theta$, and $m_v(I_t) = 1$.

From any bba $m(.)$, the belief function $Bel(.)$ and the plausibility function $Pl(.)$ are defined for $\forall X \in 2^\Theta$ as:

$$\begin{cases} Bel(X) = \sum_{Y \in 2^\Theta | Y \subseteq X} m(Y) \\ Pl(X) = \sum_{Y \in 2^\Theta | X \cap Y \neq \emptyset} m(Y) \end{cases} \tag{32}$$

$Bel(X)$ represents the whole mass of belief that comes from all subsets of Θ included in X. It is interpreted as the lower bound of the probability of X, i.e. $P_{\min}(X)$. $Bel(.)$ is a subadditive measure since $\sum_{\theta_i \in \Theta} Bel(\theta_i) \leq 1$. $Pl(X)$ represents the whole mass of belief that comes from all subsets of Θ compatible with X (i.e., those intersecting X). $Pl(X)$ is interpreted as the upper bound of the probability of X, i.e. $P_{\max}(X)$. $Pl(.)$ is a superadditive measure since $\sum_{\theta_i \in \Theta} Pl(\theta_i) \geq 1$. $Bel(X)$ and $Pl(X)$ are classically seen [4] as lower and upper bounds of an unknown probability $P(.)$, and one has the following inequality satisfied $\forall X \in 2^\Theta$: $Bel(X) \leq P(X) \leq Pl(X)$. The belief function $Bel(.)$ (and the plausibility function $Pl(.)$) built from any Bayesian bba $m(.)$ can be interpreted as a (subjective) conditional probability measure provided by a given source of evidence, because if the bba $m(.)$ is *Bayesian* the following equality always holds [4]: $Bel(X) = Pl(X) = P(X)$.

B. Dempster's rule of combination

Dempster's rule of combination, denoted DS rule[5] is a mathematical operation, represented symbolically by $\oplus$, which corresponds to the normalized conjunctive fusion rule. Based on Shafer's model of Θ, the combination of $s > 1$ independent and distinct sources of evidences characterized by their bba $m_1(.)$, ..., $m_s(.)$ related to the same frame of discernment Θ is denoted $m_{DS}(.) = [m_1 \oplus \ldots \oplus m_s](.)$. The quantity $m_{DS}(.)$ is defined mathematically as follows: $m_{DS}(\emptyset) \triangleq 0$ and $\forall X \neq \emptyset \in 2^\Theta$

$$m_{DS}(X) \triangleq \frac{m_{12\ldots s}(X)}{1 - K_{12\ldots s}} \tag{33}$$

where the conjunctive agreement on X is given by:

$$m_{12\ldots s}(X) \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = X}} m_1(X_1) m_2(X_2) \ldots m_s(X_s) \tag{34}$$

and where the global conflict is given by:

$$K_{12\ldots s} \triangleq \sum_{\substack{X_1, X_2, \ldots, X_s \in 2^\Theta \\ X_1 \cap X_2 \cap \ldots \cap X_s = \emptyset}} m_1(X_1) m_2(X_2) \ldots m_s(X_s) \tag{35}$$

When $K_{12\ldots s} = 1$, the s sources are in total conflict and their combination cannot be computed with DS rule because Eq. (33) is mathematically not defined due to 0/0 indeterminacy [4]. DS rule is commutative and associative which makes it very attractive from engineering implementation standpoint.

It has been proved in [4] that the vacuous bba $m_v(.)$ is a neutral element for DS rule because $[m \oplus m_v](.) = [m_v \oplus m](.) = m(.)$ for any bba $m(.)$ defined on 2^Θ. This property looks reasonable since a total ignorant source should not impact the fusion result because it brings no information that can be helpful for the discrimination between the elements of the power set 2^Θ.

IV. Analysis of compatibility of Dempster's rule with Bayes rule

To analyze the compatibility of Dempster's rule with Bayes rule, we need to work in the probabilistic framework because Bayes fusion rule has been developed only in this theoretical framework. So in the sequel, we will manipulate only probability mass functions (pmfs), related with Bayesian bba's in the Belief Function framework. This perfectly justifies the restriction of singleton bba as a prior bba since we want to manipulate prior probabilities to make a fair comparison of results provided by both rules. If Dempster's rule is a true (consistent) generalization of Bayes fusion rule, it must provide same results as Bayes rule when combining Bayesian bba's, otherwise Dempster's rule cannot be fairly claimed to be a generalization of Bayes fusion rule. In this section, we analyze the real (partial or total) compatibility of Dempster's rule with Bayes fusion rule. Two important cases must be analyzed depending on the nature of the prior information $P(X)$ one has in hands for performing the fusion of the sources. These

[4]The set $\{\theta_1, \theta_2, \ldots, \theta_N\}$ and the complete ignorance $\theta_1 \cup \theta_2 \cup \ldots \cup \theta_N$ are both denoted Θ in DST.

[5]We denote it DS rule because it has been proposed historically by Dempster [2], [3], and widely promoted by Shafer in the development of DST [4].

sources to combine will be characterized by the following Bayesian bba's:

$$\begin{cases} m_1(.) \triangleq \{m_1(\theta_i) = P(X = x_i|Z_1), i = 1, 2, \ldots, N\} \\ \quad \vdots \quad \vdots \qquad\qquad \vdots \\ m_s(.) \triangleq \{m_s(\theta_i) = P(X = x_i|Z_s), i = 1, 2, \ldots, N\} \end{cases} \tag{36}$$

The prior information is characterized by a given bba denoted as $m_0(.)$ that can be defined either on 2^Θ, or only on Θ if we want to deal for the needs of our analysis with a Bayesian prior. In the latter case, if $m_0(.) \triangleq \{m_0(\theta_i) = P(X = x_i), i = 1, 2, \ldots, N\}$ then $m_0(.)$ plays the same role as the prior pmf $P(X)$ in the probabilistic framework.

When considering a non vacuous prior $m_0(.) \neq m_v(.)$, we denote Dempster's combination of s sources symbolically as:

$$m_{DS}(.) = DS(m_1(.), \ldots, m_s(.); m_0(.))$$

When the prior bba is vacuous $m_0(.) = m_v(.)$ then $m_0(.)$ has no impact on Dempster's fusion result, and so we denote symbolically Dempster's rule as:

$$\begin{aligned} m_{DS}(.) &= DS(m_1(.), \ldots, m_s(.); m_v(.)) \\ &= DS(m_1(.), \ldots, m_s(.)) \end{aligned}$$

A. Case 1: Uniform Bayesian prior

It is important to note that Dempster's fusion formula proposed by Shafer in [4] and recalled in Eq. (33) makes no real distinction between the nature of sources to combine (if they are posterior or prior information). In fact, the formula (33) reduces exactly to Bayes rule given in Eq. (25) if the bba's to combine are Bayesian and if the prior information is either uniform or vacuous. Stated otherwise the following functional equality holds

$$\begin{aligned} &DS(m_1(.), \ldots, m_s(.); m_0(.)) \equiv \\ &\qquad Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X)) \end{aligned} \tag{37}$$

as soon as all bba's $m_i(.)$, $i = 1, 2, \ldots, s$ are Bayesian and coincide with $P(X|Z_i)$, $P(X)$ is uniform, and either the prior bba $m_0(.)$ is vacuous ($m_0(.) = m_v(.)$), or $m_0(.)$ is the uniform Bayesian bba.

Example 3: Let us consider $\Theta(X) = \{x_1, x_2, x_3\}$ with two distinct sources providing the following Bayesian bba's

$$\begin{cases} m_1(x_1) = P(X = x_1|Z_1) = 0.2 \\ m_1(x_2) = P(X = x_2|Z_1) = 0.3 \\ m_1(x_3) = P(X = x_3|Z_1) = 0.5 \end{cases} \quad \text{and} \quad \begin{cases} m_2(x_1) = 0.5 \\ m_2(x_2) = 0.1 \\ m_2(x_3) = 0.4 \end{cases}$$

• If we choose as prior $m_0(.)$ the vacuous bba, that is $m_0(x_1 \cup x_2 \cup x_3) = 1$, then one will get

$$\begin{cases} m_{DS}(x_1) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_1)m_2(x_1)m_0(x_1 \cup x_2 \cup x_3) \\ &= \frac{1}{1-0.67} 0.2 \cdot 0.5 \cdot 1 = \frac{0.10}{0.33} \approx 0.3030 \\ m_{DS}(x_2) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_2)m_2(x_2)m_0(x_1 \cup x_2 \cup x_3) \\ &= \frac{1}{1-0.67} 0.3 \cdot 0.1 \cdot 1 = \frac{0.03}{0.33} \approx 0.0909 \\ m_{DS}(x_3) &= \frac{1}{1-K_{12}^{vacuous}} m_1(x_3)m_2(x_3)m_0(x_1 \cup x_2 \cup x_3) \\ &= \frac{1}{1-0.67} 0.5 \cdot 0.4 \cdot 1 = \frac{0.20}{0.33} \approx 0.6061 \end{cases}$$

with

$$\begin{aligned} K_{12}^{vacuous} = 1 &- m_1(x_1)m_2(x_1)m_0(x_1 \cup x_2 \cup x_3) \\ &- m_1(x_2)m_2(x_2)m_0(x_1 \cup x_2 \cup x_3) \\ &- m_1(x_3)m_2(x_3)m_0(x_1 \cup x_2 \cup x_3) = 0.67 \end{aligned}$$

• If we choose as prior $m_0(.)$ the uniform Bayesian bba given by $m_0(x_1) = m_0(x_2) = m_0(x_3) = 1/3$, then we get

$$\begin{cases} m_{DS}(x_1) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_1)m_2(x_1)m_0(x_1) \\ &= \frac{1}{1-0.89} 0.2 \cdot 0.5 \cdot 1/3 = \frac{0.10/3}{0.11} \approx 0.3030 \\ m_{DS}(x_2) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_2)m_2(x_2)m_0(x_2) \\ &= \frac{1}{1-0.89} 0.3 \cdot 0.1 \cdot 1/3 = \frac{0.03/3}{0.11} \approx 0.0909 \\ m_{DS}(x_3) &= \frac{1}{1-K_{12}^{uniform}} m_1(x_3)m_2(x_3)m_0(x_3) \\ &= \frac{1}{1-0.89} 0.5 \cdot 0.4 \cdot 1/3 = \frac{0.20/3}{0.11} \approx 0.6061 \end{cases}$$

where the degree of conflict when $m_0(.)$ is Bayesian and uniform is now given by $K_{12}^{uniform} = 0.89$.

Clearly $K_{12}^{uniform} \neq K_{12}^{vacuous}$, but the fusion results obtained with two distinct priors $m_0(.)$ (vacuous or uniform) are the same because of the algebraic simplification by 1/3 in Dempster's fusion formula when using uniform Bayesian bba. When combining Bayesian bba's $m_1(.)$ and $m_2(.)$, the vacuous prior and uniform prior $m_0(.)$ have therefore no impact on the result. Indeed, they contain no information that may help to prefer one particular state x_i with respect to the other ones, even if the level of conflict is different in both cases. So, the level of conflict doesn't matter at all in such Bayesian case. As already stated, what really matters is only the distribution of relative agreement factors. It can be easily verified that we obtain same results when applying Bayes Eq. (14), or (16).

Only in such very particular cases (i.e. Bayesian bba's, and vacuous or Bayesian uniform priors), Dempster's rule is fully consistent with Bayes fusion rule. So the claim that Dempster's is a generalization of Bayes rule is true in this very particular case only, and that is why such claim has been widely used to defend Dempster's rule and DST thanks to its compatibility with Bayes fusion rule in that very particular case. Unfortunately, such compatibility is only partial and not general because it is not longer valid when considering the more general cases involving non uniform Bayesian prior bba's as shown in the next subsection.

B. Case 2: Non uniform Bayesian prior

Let us consider Dempster's fusion of Bayesian bba's with a Bayesian non uniform prior $m_0(.)$. In such case it is easy to check from the general structures of Bayes fusion rule (16) and Dempster's fusion rule (33) that these two rules are incompatible. Indeed, in Bayes rule one divides each posterior source $m_i(x_j)$ by $\sqrt[s]{m_0(x_j)}$, $i = 1, 2, \ldots s$, whereas the prior source $m_0(.)$ is combined in a pure conjunctive manner by Dempster's rule with the bba's $m_i(.)$, $i = 1, 2, \ldots s$, as if $m_0(.)$ was a simple additional source. This difference of processing prior information between the two approaches explains clearly the incompatibility of Dempster's rule with Bayes rule when Bayesian prior bba is not uniform. This incompatibility is illustrated in the next simple example. Mahler and Fixsen have already proposed in [23], [24], [25] a modification of

Dempster's rule to force it to be compatible with Bayes rule when combining Bayesian bba's. The analysis of such modified Dempster's rule is out of the scope of this paper.

Example 4: Let us consider the same frame $\Theta(X)$, and same bba's $m_1(.)$ and $m_2(.)$ as in the Example 3. Suppose that the prior information is Bayesian and non uniform as follows: $m_0(x_1) = P(X = x_1) = 0.6$, $m_0(x_2) = P(X = x_2) = 0.3$ and $m_0(x_3) = P(X = x_3) = 0.1$. Applying Bayes rule (12) yields:

$$\begin{cases} P(x_1|Z_1 \cap Z_2) &= \frac{A_2(x_1)}{GA_2} = \frac{0.2 \cdot 0.5/0.6}{2.2667} = \frac{0.1667}{2.2667} \approx 0.0735 \\ P(x_2|Z_1 \cap Z_2) &= \frac{A_2(x_2)}{GA_2} = \frac{0.3 \cdot 0.1/0.3}{2.2667} = \frac{0.1000}{2.2667} \approx 0.0441 \\ P(x_3|Z_1 \cap Z_2) &= \frac{A_2(x_3)}{GA_2} = \frac{0.5 \cdot 0.4/0.1}{2.2667} = \frac{2.0000}{2.2667} \approx 0.8824 \end{cases}$$

Applying Dempster's rule yields $m_{DS}(x_i) \neq P(x_i|Z_1 \cap Z_2)$ because:

$$\begin{cases} m_{DS}(x_1) &= \frac{1}{1-0.9110} \cdot 0.2 \cdot 0.5 \cdot 0.6 = \frac{0.060}{0.089} \approx 0.6742 \\ m_{DS}(x_2) &= \frac{1}{1-0.9110} \cdot 0.3 \cdot 0.1 \cdot 0.3 = \frac{0.009}{0.089} \approx 0.1011 \\ m_{DS}(x_3) &= \frac{1}{1-0.9110} \cdot 0.5 \cdot 0.4 \cdot 0.1 = \frac{0.020}{0.089} \approx 0.2247 \end{cases}$$

Therefore, one has in general[6]:

$$DS(m_1(.), \ldots, m_s(.); m_0(.)) \neq Bayes(P(X|Z_1), \ldots, P(X|Z_s); P(X)) \quad (38)$$

V. Conclusions

In this paper, we have analyzed in details the expression and the properties of Bayes rule of combination based on statistical conditional independence assumption, as well as the emblematic Dempster's rule of combination of belief functions introduced by Shafer in his Mathematical Theory of evidence. We have clearly explained from a theoretical standpoint, and also on simple examples, why Dempster's rule is not a generalization of Bayes rule in general. The incompatibility of Dempster's rule with Bayes rule is due to its impossibility to deal with non uniform Bayesian priors in the same manner as Bayes rule does. Dempster's rule turns to be compatible with Bayes rule only in two very particular cases: 1) if all the Bayesian bba's to combine (including the prior) focus on same state (i.e. there is a perfect conjunctive consensus between the sources), or 2) if all the bba's to combine (excluding the prior) are Bayesian, and if the prior bba cannot help to discriminate a particular state of the frame of discernment (i.e. the prior bba is either vacuous, or Bayesian and uniform). Except in these two very particular cases, Dempster's rule is totally incompatible with Bayes rule. Therefore, Dempster's rule cannot be claimed to be a generalization of Bayes fusion rule, even when the bba's to combine are Bayesian.

Acknowledgment

This study was co-supported by Grant for State Key Program for Basic Research of China (973) (No. 2013CB329405), National NSF of China (No.61104214, No. 61203222), and also partly supported by the project AComIn, grant 316087, funded by the FP7 Capacity Programme.

[6]but in the very degenerate case when manipulating *deterministic* Bayesian bba's, which is of little practical interest from the fusion standpoint.

References

[1] L.A. Zadeh, On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, CA, U.S.A., 1979.

[2] A. Dempster, Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Statist., Vol. 38, pp. 325–339, 1967.

[3] A. Dempster, A generalization of bayesian inference, J. R. Stat. Soc. B 30, pp. 205–247, 1968.

[4] G. Shafer, A Mathematical theory of evidence, Princeton University Press, Princeton, NJ, U.S.A., 1976.

[5] L.A. Zadeh, Book review: A mathematical theory of evidence, The Al Magazine, Vol. 5, No. 3, pp. 81-83, 1984.

[6] L.A. Zadeh, A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The Al Magazine, Vol. 7, No. 2, pp. 85–90, 1986.

[7] J. Lemmer, Confidence factors, empiricism and the Dempster-Shafer theory of evidence, Proc. of UAI '85, pp. 160–176, Los Angeles, CA, U.S.A., July 10–12, 1985.

[8] J. Pearl, Do we need higher-order probabilities, and, if so, what do they mean?, Proc. of UAI '87, pp. 47–60, Seattle, WA, U.S.A., July 10–12, 1987.

[9] F. Voorbraak, On the justification of Demspster's rule of combination, Dept. of Phil., Utrecht Univ., Logic Group, Preprint Ser., No. 42, Dec. 1988.

[10] G. Shafer, Perspectives on the theory and practice of belief functions, IJAR, Vol. 4, No. 5–6, pp. 323–362, 1990.

[11] J. Pearl, Reasoning with belief functions: an analysis of compatibility, IJAR, Vol. 4, No. 5–6, pp. 363–390, 1990.

[12] J. Pearl, Rejoinder of comments on "Reasoning with belief functions: An analysis of compatibility", IJAR, Vol. 6, No. 3, pp. 425–443, May 1992.

[13] G.M. Provan, The validity of Dempster-Shafer belief functions, IJAR, Vol. 6., No. 3, pp. 389–399, May 1992.

[14] P. Wang, A defect in Dempster-Shafer theory, Proc. of UAI '94, pp. 560–566, Seattle, WA, U.S.A., July 29–31, 1994.

[15] A. Gelman, The boxer, the wrestler, and the coin flip: a paradox of robust bayesian inference and belief functions, American Statistician, Vol. 60, No. 2, pp. 146–150, 2006.

[16] F. Smarandache, J. Dezert (Editors), Applications and advances of DSmT for information fusion, Vol. 3, ARP, U.S.A., 2009. http://fs.gallup.unm.edu/DSmT.htm

[17] J. Dezert, P. Wang, A. Tchamova, On the validity of Dempster-Shafer theory, Proc. of Fusion 2012 Int. Conf., Singapore, July 9–12, 2012.

[18] A. Tchamova, J. Dezert, On the Behavior of Dempster's rule of combination and the foundations of Dempster-Shafer theory, (best paper award), Proc. of 6th IEEE Int. Conf. on Intelligent Systems IS '12, Sofia, Bulgaria, Sept. 6–8, 2012.

[19] R.P. Mahler, Statistical multisource-multitarget information fusion, Chapter 4, Artech House, 2007.

[20] X.R. Li, Probability, random signals and statistics, CRC Press, 1999.

[21] P.E. Pfeiffer, Applied probability, Connexions Web site, Aug. 31, 2009. http://cnx.org/content/col10708/1.6/

[22] G. Shafer, Non-additive probabilities in the work of Bernoulli and Lambert, in Archive for History of Exact Sciences, C. Truesdell (Ed.), Springer-Verlag, Berlin, Vol. 19, No. 4, pp. 309–370, 1978.

[23] R.P. Mahler, Using a priori evidence to customize Dempster-Shafer theory, Proc. of 6th Nat. Symp. on Sensor Fusion, Vol. 1, pp. 331–345, Orlando, FL, U.S.A., April 13–15, 1993.

[24] R.P. Mahler, Classification when a priori evidence is ambiguous, Proc. SPIE Conf. on Opt. Eng. in Aerospace Sensing (Automatic Object Recognition IV), pp. 296–304, Orlando, FL, U.S.A., April 4–8, 1994.

[25] D. Fixsen, R.P. Mahler, The modified Dempster-Shafer approach to classification, IEEE Trans. on SMC, Part A, Vol. 27, No. 1, pp. 96–104, 1997.

On the Consistency of PCR6 with the Averaging Rule and its Application to Probability Estimation

Florentin Smarandache
Jean Dezert

Originally published as Smarandache F., Dezert J., *On the consistency of PCR6 with the averaging rule and its application to probability estimation*, Proc. of Fusion 2013 Int. Conference on Information Fusion, Istanbul, Turkey, July 9-12, 2013, and reprinted with permission. (with typos corrections).

***Abstract*—Since the development of belief function theory introduced by Shafer in seventies, many combination rules have been proposed in the literature to combine belief functions specially (but not only) in high conflicting situations because the emblematic Dempster's rule generates counter-intuitive and unacceptable results in practical applications. Many attempts have been done during last thirty years to propose better rules of combination based on different frameworks and justifications. Recently in the DSmT (Dezert-Smarandache Theory) framework, two interesting and sophisticate rules (PCR5 and PCR6 rules) have been proposed based on the Proportional Conflict Redistribution (PCR) principle. These two rules coincide for the combination of two basic belief assignments, but they differ in general as soon as three or more sources have to be combined altogether because the PCR used in PCR5 and in PCR6 are different. In this paper we show why PCR6 is better than PCR5 to combine three or more sources of evidence and we prove the coherence of PCR6 with the simple Averaging Rule used classically to estimate the probability based on the frequentist interpretation of the probability measure. We show that such probability estimate cannot be obtained using Dempster-Shafer (DS) rule, nor PCR5 rule.**

Keywords: Information fusion, belief functions, PCR6, PCR5, DSmT, frequentist probability.

I. Introduction

In this paper, we work with belief functions [1] defined from the finite and discrete frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$. In Dempster-Shafer Theory (DST) framework, basic belief assignments (bba's) provided by the distinct sources of evidence are defined on the fusion space $2^\Theta = (\Theta, \cup)$ consisting in the power-set of Θ, that is the set of elements of Θ and those generated from Θ with the union set operator. Such fusion space assumes that the elements of Θ are non-empty, exhaustive and exclusive, which is called Shafer's model of Θ. More generally, in Dezert-Smarandache Theory (DSmT) [2], the fusion space denoted G^Θ can also be either the hyper-power set $D^\Theta = (\Theta, \cup, \cap)$ (Dedekind's lattice), or super-power set[1] $S^\Theta = (\Theta, \cup, \cap, c(.))$ depending on the underlying model of the frame of discernment we choose to fit with the nature of the problem. Details on DSm models are given in [2], Vol. 1.

We assume that $s \geq 2$ basic belief assignments (bba's) $m_i(.)$, $i = 1, 2, \ldots, s$ provided by s distinct sources of evidences defined on the fusion space G^Θ are available and we need to combine them for a final decision-making purpose. For doing this, many rules of combination have been proposed in the literature, the most emblematic ones being the simple Averaging Rule, Dempster-Shafer (DS) rule, and more recently the PCR5 and PCR6 fusion rules.

The contribution of this paper is to analyze in deep the behavior of PCR5 and PCR6 fusion rules and to explain why we consider more preferable to use PCR6 rule rather than PCR5 rule for combining several distinct sources of evidence altogether. We will show in details the strong relationship between PCR6 and the averaging fusion rule which is commonly used to estimate the probabilities in the classical frequentist interpretation of probabilities.

This paper is organized as follows. In section II, we briefly recall the background on belief functions and the main fusion rules used in this paper. Section III demonstrates the consistency of PCR6 fusion rule with the Averaging Rule for binary masses in total conflict as well as the ability of PCR6 to discriminate asymmetric fusion cases for the fusion of Bayesian bba's. Section IV shows that PCR6 can also be used to estimate empirical probability in a simple (coin tossing) random experiment. Section V will conclude and open challenging problem about the recursivity of fusion rules formulas that are sought for efficient implementations.

II. Background on belief functions

A. Basic belief assignment

Lets' consider a finite discrete frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$, $n > 1$ of the fusion problem under consideration and its fusion space G^Θ which can be chosen either as 2^Θ, D^Θ or S^Θ depending on the model that fits with the problem. A basic belief assignment (bba) associated with a given source of evidence is defined as the mapping $m(.) : G^\Theta \to [0, 1]$ satisfying $m(\emptyset) = 0$ and $\sum_{A \in G^\Theta} m(A) = 1$. The quantity $m(A)$ is called mass of belief of A committed by the source of evidence. If $m(A) > 0$ then A is called a focal element of the bba $m(.)$. When all focal elements are singletons and $G^\Theta = 2^\Theta$ then $m(.)$ is called a Bayesian bba [1] and it is homogeneous to a (possibly subjective) probability measure. The vacuous bba representing a totally ignorant source is defined as $m_v(\Theta) = 1$. Belief and plausibility functions are defined by

$$\text{Bel}(A) = \sum_{\substack{B \subseteq A \\ B \in G^\Theta}} m(B) \quad \text{and} \quad \text{Pl}(A) = \sum_{\substack{B \cap A \neq \emptyset \\ B \in G^\Theta}} m(B) \qquad (1)$$

[1] $\cap$ and $c(.)$ are respectively the set intersection and complement operators.

B. Fusion rules

The main information fusion problem in the belief function frameworks (DST or DSmT) is how to combine efficiently several distinct sources of evidence represented by $m_1(.)$, $m_2(.)$, ..., $m_s(.)$ ($s \geq 2$) bba's defined on G^Θ. Many rules have been proposed for such task – see [2], Vol. 2, for a detailed list of fusion rules – and we focus here on the following ones: 1) the Averaging Rule because it is the simplest one and it is used to empirically estimate probabilities in random experiment, 2) DS rule because it was historically proposed in DST, and 3) PCR5 and PCR6 rules because they were proposed in DSmT and have shown to provide better results than the DS rule in all applications where they have been tested so far. So we just briefly recall how these rules are mathematically defined.

• Averaging fusion rule $m_{1,2,...,s}^{Average}(.)$

For any X in G^Θ, $m_{1,2,...,s}^{Average}(X)$ is defined by

$$m_{1,2,...,s}^{Average}(X) = Average(m_1, m_2, \ldots, m_s) \triangleq \frac{1}{s}\sum_{i=1}^{s} m_i(X) \quad (2)$$

Note that the vacuous bba $m_v(\Theta) = 1$ is not a neutral element for this rule. This Averaging Rule is commutative but it is not associative because in general

$$m_{1,2,3}^{Average}(X) = \frac{1}{3}[m_1(X) + m_2(X) + m_3(X)]$$

is different from

$$m_{(1,2),3}^{Average}(X) = \frac{1}{2}[\frac{m_1(X) + m_2(X)}{2} + m_3(X)]$$

which is also different from

$$m_{1,(2,3)}^{Average}(X) = \frac{1}{2}[m_1(A) + \frac{m_2(X) + m_3(X)}{2}]$$

and also from

$$m_{(1,3),2}^{Average}(X) = \frac{1}{2}[\frac{m_1(X) + m_3(X)}{2} + m_2(X)]$$

In fact, it is easy to prove that the following recursive formula holds

$$m_{1,2,...,s}^{Average}(X) = \frac{s-1}{s} m_{1,2,...,s-1}^{Average}(X) + \frac{1}{s} m_s(X) \quad (3)$$

This simple averaging fusion rule has been used since more than two centuries for estimating empirically the probability measure in random experiments [3], [4].

• Dempster-Shafer fusion rule $m_{1,2,...,s}^{DS}(.)$

In DST framework, the fusion space G^Θ equals the power-set 2^Θ because Shafer's model of the frame Θ is assumed. The combination of $s \geq 2$ distinct sources of evidences characterized by the bba's $m_i(.)$, $i = 1, 2, \ldots, s$, is done with DS rule as follows [1]: $m_{1,2,...,s}^{DS}(\emptyset) = 0$ and for all $X \neq \emptyset$ in 2^Θ

$$m_{1,2,...,s}^{DS}(X) \triangleq \frac{1}{K_{1,2,...,s}} \sum_{\substack{X_1, X_2, ..., X_s \in 2^\Theta \\ X_1 \cap X_2 \cap ... \cap X_s = X}} \prod_{i=1}^{s} m_i(X_i) \quad (4)$$

where the numerator of (4) is the mass of belief on the conjunctive consensus on X, and where $K_{1,2,...,s}$ is a normalization constant defined by

$$K_{1,2,...,s} = \sum_{\substack{X_1, X_2, ..., X_s \in 2^\Theta \\ X_1 \cap X_2 \cap ... \cap X_s \neq \emptyset}} \prod_{i=1}^{s} m_i(X_i) = 1 - m_{1,2,...,s}(\emptyset)$$

The total degree of conflict between the s sources of evidences is defined by

$$m_{1,2,...,s}(\emptyset) = \sum_{\substack{X_1, X_2, ..., X_s \in 2^\Theta \\ X_1 \cap X_2 \cap ... \cap X_s = \emptyset}} \prod_{i=1}^{s} m_i(X_i)$$

The sources are said in total conflict when $m_{1,2,...,s}(\emptyset) = 1$.

The vacuous bba $m_v(\Theta) = 1$ is a neutral element for DS rule and DS rule is commutative and associative. It remains the milestone fusion rule of DST. The doubts on the validity of such fusion rule has been discussed by Zadeh in 1979 [5]–[7] based on a very simple example with two highly conflicting sources of evidence. Since 1980's, many criticisms have been done about the behavior and justification of such DS rule. More recently, Dezert et al. in [8], [9] have put in light other counter-intuitive behaviors of DS rule even in low conflicting cases and showed serious flaws in logical foundations of DST.

• PCR5 and PCR6 fusion rules

To work in general fusion spaces G^Θ and to provide better fusion results in all (low or high conflicting) situations, several fusion rules have been developed in DSmT framework [2]. Among them, two fusion rules called PCR5 and PCR6 based on the proportional conflict redistribution (PCR) principle have been proved to work efficiently in all different applications where they have been used so far. The PCR principle transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved.

The general principle of PCR consists:

1) to apply the conjunctive rule;
2) calculate the total or partial conflicting masses;
3) then redistribute the (total or partial) conflicting mass proportionally on non-empty sets according to the integrity constraints one has for the frame Θ.

Because the proportional transfer can be done in two different ways, this has yielded to two different fusion rules. The PCR5 fusion rule has been proposed by Smarandache and Dezert in [2], Vol. 2, Chap. 1, and PCR6 fusion rule has been proposed by Martin and Osswald in [2], Vol. 2, Chap. 2.

We will not present in deep these two fusion rules since they have already been discussed in details with many examples in the aforementioned references but we only give their expressions for convenience here.

The general formula of PCR5 for the combination of $s \geq 2$ sources is given by $m_{1,2,...,s}^{PCR5}(\emptyset) = 0$ and for $X \neq \emptyset$ in G^Θ

$$m_{1,2,...,s}^{PCR5}(X) = m_{1,2,...,s}(X) + \sum_{\substack{2\leq t\leq s \\ 1\leq r_1,...,r_t\leq s \\ 1\leq r_1<r_2<...<r_{t-1}<(r_t=s)}} \sum_{\substack{X_{j_2},...,X_{j_t}\in G^\Theta\setminus\{X\} \\ \{j_2,...,j_t\}\in\mathcal{P}^{t-1}(\{1,...,n\}) \\ X\cap X_{j_2}\cap...\cap X_{j_s}=\emptyset \\ \{i_1,...,i_s\}\in\mathcal{P}^s(\{1,...,s\})}} \frac{(\prod_{k_1=1}^{r_1} m_{i_{k_1}}(X)^2)\cdot[\prod_{l=2}^{t}(\prod_{k_l=r_{l-1}+1}^{r_l} m_{i_{k_l}}(X_{j_l})]}{(\prod_{k_1=1}^{r_1} m_{i_{k_1}}(X)) + [\sum_{l=2}^{t}(\prod_{k_l=r_{l-1}+1}^{r_l} m_{i_{k_l}}(X_{j_l})]} \quad (5)$$

where i, j, k, r, s and t in (5) are integers. $m_{1,2,...,s}(X)$ corresponds to the conjunctive consensus on X between s sources and where all denominators are different from zero. If a denominator is zero, that fraction is discarded; $\mathcal{P}^k(\{1,2,\ldots,n\})$ is the set of all subsets of k elements from $\{1,2,\ldots,n\}$ (permutations of n elements taken by k), the order of elements doesn't count.

The general formula of PCR6 proposed by Martin and Osswald for the combination of $s \geq 2$ sources is given by $m_{1,2,...,s}^{PCR6}(\emptyset) = 0$ and for $X \neq \emptyset$ in G^Θ

$$m_{1,2,...,s}^{PCR6}(X) = m_{1,2,...,s}(X) + \sum_{i=1}^{s} m_i(X)^2 \sum_{\substack{\bigcap_{k=1}^{s-1} Y_{\sigma_i(k)}\cap X\equiv\emptyset \\ (Y_{\sigma_i(1)},...,Y_{\sigma_i(s-1)})\in(G^\Theta)^{s-1}}} \left(\frac{\prod_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)})}{m_i(X)+\sum_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)})} \right) \quad (6)$$

where σ_i counts from 1 to s avoiding i:

$$\begin{cases} \sigma_i(j) = j & \text{if } j < i, \\ \sigma_i(j) = j+1 & \text{if } j \geq i, \end{cases} \quad (7)$$

Since Y_i is a focal element of expert/source i, $m_i(X)+\sum_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)}) \neq 0$.

The general PCR5 and PCR6 formulas (5)–(6) are rather complicate and not very easy to understand. From the implementation point of view, PCR6 is much simple to implement than PCR5. For convenience, very basic (not optimized) Matlab codes of PCR5 and PCR6 fusion rules can be found in [2], [10] and from the toolboxes repository on the web [11]. The PCR5 and PCR6 fusion rules are commutative but not associative, like the averaging fusion rule, but the vacuous belief assignment is a neutral element for these PCR fusion rules.

The PCR5 and PCR6 fusion rules simplify greatly and coincide for the combination of two sources ($s = 2$). In such simplest case, one always gets the resulting bba $m_{PCR5/6}(.) = m_{1,2}^{PCR6}(.) = m_{1,2}^{PCR5}(.)$ expressed as $m_{PCR5/6}(\emptyset) = 0$ and for all $X \neq \emptyset$ in G^Θ

$$m_{PCR5/6}(X) = \sum_{\substack{X_1,X_2\in G^\Theta \\ X_1\cap X_2=X}} m_1(X_1)m_2(X_2) + \sum_{\substack{Y\in G^\Theta\setminus\{X\} \\ X\cap Y=\emptyset}} [\frac{m_1(X)^2 m_2(Y)}{m_1(X)+m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X)+m_1(Y)}] \quad (8)$$

where all denominators in (8) are different from zero. If a denominator is zero, that fraction is discarded. All propositions/sets are in a canonical form.

Example 1: See [2], Vol.2, Chap. 1 for more examples.

Let's consider the frame of discernment $\Theta = \{A, B\}$ of exclusive elements. Here Shafer's model holds so that $G^\Theta = 2^\Theta = \{\emptyset, A, B, A\cup B\}$. We consider two sources of evidences providing the following bba's

$$m_1(A) = 0.6 \quad m_1(B) = 0.3 \quad m_1(A\cup B) = 0.1$$

$$m_2(A) = 0.2 \quad m_2(B) = 0.3 \quad m_2(A\cup B) = 0.5$$

Then the conjunctive consensus yields :

$$m_{12}(A) = 0.44 \quad m_{12}(B) = 0.27 \quad m_{12}(A\cup B) = 0.05$$

with the conflicting mass

$$\begin{aligned} m_{12}(A\cap B = \emptyset) &= m_1(A)m_2(B) + m_1(B)m_2(A) \\ &= 0.18 + 0.06 = 0.24 \end{aligned}$$

One sees that only A and B are involved in the derivation of the conflicting mass, but not $A \cup B$. With PCR5/6, one redistributes the partial conflicting mass 0.18 to A and B proportionally with the masses $m_1(A)$ and $m_2(B)$ assigned to A and B respectively, and also the partial conflicting mass 0.06 to A and B proportionally with the masses $m_2(A)$ and $m_1(B)$ assigned to A and B respectively, thus one gets two weighting factors of the redistribution for each corresponding set A and B respectively. Let x_1 be the conflicting mass to be redistributed to A, and y_1 the conflicting mass redistributed to B from the first partial conflicting mass 0.18. This first partial proportional redistribution is then done according

$$\frac{x_1}{0.6} = \frac{y_1}{0.3} = \frac{x_1+y_1}{0.6+0.3} = \frac{0.18}{0.9} = 0.2$$

whence $x_1 = 0.6\cdot 0.2 = 0.12$, $y_1 = 0.3\cdot 0.2 = 0.06$. Now let x_2 be the conflicting mass to be redistributed to A, and y_2 the conflicting mass redistributed to B from the second the partial conflicting mass 0.06. This second partial proportional redistribution is then done according

$$\frac{x_2}{0.2} = \frac{y_2}{0.3} = \frac{x_2+y_2}{0.2+0.3} = \frac{0.06}{0.5} = 0.12$$

whence $x_2 = 0.2 \cdot 0.12 = 0.024$, $y_2 = 0.3 \cdot 0.12 = 0.036$. Thus one finally gets:

$$\begin{aligned} m_{PCR5/6}(A) &= 0.44 + 0.12 + 0.024 = 0.584 \\ m_{PCR5/6}(B) &= 0.27 + 0.06 + 0.036 = 0.366 \\ m_{PCR5/6}(A \cup B) &= 0.05 + 0 = 0.05 \end{aligned}$$

• The difference between PCR5 and PCR6 fusion rules

For the two sources case, PCR5 and PCR6 fusion rules coincide. As soon as three (or more) sources are involved in the fusion process, PCR5 and PCR6 differ in the way the proportional conflict redistribution is done. For example, let's consider three sources with bba's $m_1(.)$, $m_2(.)$ and $m_3(.)$, $A \cap B = \emptyset$ for the model of the frame Θ, and $m_1(A) = 0.6$, $m_2(B) = 0.3$, $m_3(B) = 0.1$.

– With PCR5, the mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ corresponding to a conflict is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR5} = 0.01714$ and $x_B^{PCR5} = 0.00086$ because the proportionalization requires

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B)m_3(B)}$$

that is

$$\frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{0.03} = \frac{0.018}{0.6 + 0.03} \approx 0.02857$$

Thus

$$\begin{cases} x_A^{PCR5} = 0.60 \cdot 0.02857 \approx 0.01714 \\ x_B^{PCR5} = 0.03 \cdot 0.02857 \approx 0.00086 \end{cases}$$

– With the PCR6 fusion rule, the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR6} = 0.0108$ and $x_B^{PCR6} = 0.0072$ because the PCR6 proportionalization is done as follows:

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_B^{PCR6}}{m_2(B) + m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + (m_2(B) + m_3(B))}$$

that is

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_B^{PCR6}}{0.3 + 0.1} = \frac{0.018}{0.6 + (0.3 + 0.1)} = 0.018$$

and therefore with PCR6, one gets finally the following redistributions to A and B:

$$\begin{cases} x_A^{PCR6} = 0.6 \cdot 0.018 = 0.0108 \\ x_B^{PCR6} = (0.3 + 0.1) \cdot 0.018 = 0.0072 \end{cases}$$

In [2], Vol. 2, Chap. 2, Martin and Osswald have proposed PCR6 based on intuitive considerations and the authors have shown through simulations that PCR6 is more stable than PCR5 in term of decision for combining $s > 2$ sources of evidence. Based on these results and the relative "simplicity" of implementation of PCR6 over PCR5, PCR6 has been considered more interesting/efficient than PCR5 for combining 3 (or more) sources of evidences.

III. Consistency of PCR6 with the Averaging Rule

In this section we show why we also consider PCR6 as better than PCR5 for combining bba's. But here, our argumentation is not based on particular simulation results and decision-making as done by Martin and Osswald, but on a theoretical analysis of the structure of PCR6 fusion rule itself. In particular, we show the full consistency of PCR6 rule with the averaging fusion rule used to empirically estimate probabilities in random experiments. For doing this, it is necessary to simplify the original PCR6 fusion formula (6). Such simplification has already been proposed in [12] and the PCR6 fusion rule can be in fact rewritten as

$$\begin{aligned} m_{1,2,\ldots,s}^{PCR6}(X) &= m_{1,2,\ldots,s}(X) + \\ &\sum_{k=1}^{s-1} \sum_{\substack{X_{i_1}, X_{i_2}, \ldots, X_{i_k} \in G^\Theta \setminus X \\ (\bigcap_{j=1}^{k} X_{i_j}) \cap X = \emptyset}} \sum_{(i_1, i_2, \ldots, i_k) \in \mathcal{P}^s(\{1,\ldots,s\})} \\ &[m_{i_1}(X) + m_{i_2}(X) + \ldots + m_{i_k}(X)] \cdot \\ &\cdot \frac{m_{i_1}(X) \ldots m_{i_k}(X) m_{i_{k+1}}(X_{i_{k+1}}) \ldots m_{i_s}(X_{i_s})}{m_{i_1}(X) + \ldots + m_{i_k}(X) + m_{i_{k+1}}(X_{i_{k+1}}) + \ldots + m_{i_s}(X_{i_s})} \end{aligned} \quad (9)$$

where $\mathcal{P}^s(\{1, \ldots, s\})$ is the set of all permutations of the elements $\{1, 2, \ldots, s\}$. It should be observed that X_{i_1}, X_{i_2},...,X_{i_s} may be different from each other, or some of them equal and others different, etc.

We wrote this PCR6 general formula (9) in the style of PCR5, different from Arnaud Martin & Christophe Oswald's notations, but actually doing the same thing. In order not to complicate the formula of PCR6, we did not use more summations or products after the third Sigma.

We now are able to establish the consistency of general PCR6 formula with the Averaging fusion rule for the case of binary bba's through the following theorem 1.

Theorem 1: When $s \geq 2$ sources of evidences provide binary bba's on G^Θ whose total conflicting mass is 1, then the PCR6 fusion rule coincides with the averaging fusion rule. Otherwise, PCR6 and the averaging fusion rule provide in general different results.

Proof 1: All $s \geq 2$ bba's are assumed binary, i.e. $m(X) = 0$ or 1 (two numerical values 0 and 1 only are allowed) for any bba $m(.)$ and for any set X in the focal elements. A focal element in this case is an element X such that at least one of the s binary sources assigns a mass equals to 1 to X. Let's suppose the focal elements are F_1, F_2,..., F_n.. Then the set of bba's to combine can be expressed as in the Table I. where

Table I. List of bba's to combine.

bba's \ Focal elem.	F_1	F_2	...	F_n
$m_1(.)$	$\star$	$\star$	...	$\star$
$m_2(.)$	$\star$	$\star$	...	$\star$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$m_s(.)$	$\star$	$\star$	...	$\star$

- all $\star$ are 0's or 1's;
- on each row there is only a 1 (since the sum of all masses of a bba is equal to 1) and all the other elements are 0's;

- also each column has at least an 1 (since all elements are focals; and if there was a column corresponding for example to the set F_p having only 0's, then it would result that the set F_p is not focal, i.e. that all $m(F_p) = 0$).

Using PCR6, we first need to apply the conjunctive rule to all s sources, and the result is a sum of products of the form $m_1(X_1)m_2(X_2)\dots m_s(X_s)$ where X_1, X_2,...,X_s, are the focal elements F_1, F_2,...,F_n in various permutations, with $s \geq n$. If $s > n$ some focal elements X_i are repeated in the product $m_1(X_1)m_2(X_2)\dots m_s(X_s)$. But there is only one product of the form $m_1(X_1)m_2(X_2)\dots m_s(X_s) = 1$ which is not equal to zero, i.e. that product which has each factor equals to "1" (i.e. the product that collects from each row the existing single 1). Since the total conflicting mass is equal to 1, it means that this product represents the total conflict. In this case the PCR6 formula (9) becomes

$$m_{1,2,\dots,s}^{PCR6}(X) = 0+ \sum_{k=1}^{s-1} \sum_{\substack{X_{i_1},X_{i_2},\dots,X_{i_k} \in G^\Theta \setminus X \\ (\bigcap_{j=1}^{k} X_{i_j}) \cap X = \emptyset}} \sum_{(i_1,i_2,\dots,i_k) \in \mathcal{P}^s(\{1,\dots,s\})} [1+1+\dots+1] \cdot \frac{1 \cdot 1 \cdot \dots \cdot 1 \cdot 1 \cdot \dots \cdot 1}{1+1+\dots+1+1+\dots+1} \quad (10)$$

The previous expression can be rewritten as

$$m_{1,2,\dots,s}^{PCR6}(X) = \sum_{k=1}^{s-1} \sum_{\substack{X_{i_1},X_{i_2},\dots,X_{i_k} \in G^\Theta \setminus X \\ (\bigcap_{j=1}^{k} X_{i_j}) \cap X = \emptyset}} \sum_{\substack{(i_1,i_2,\dots,i_k) \\ \in \mathcal{P}^s(\{1,\dots,s\})}} k \cdot \frac{1}{s}$$

which is equal to k/s since there is only one possible non-null product of the form $m_1(X_1)m_2(X_2)\dots m_s(X_s)$, and all other products are equal to zero. Therefore, we finally get:

$$m_{1,2,\dots,s}^{PCR6}(X) = \frac{k}{s} \quad (11)$$

where "k" is the number of bba's $m(.)$ which give $m(X) = 1$. Therefore PCR6 in this case reduces to the average of masses, which completes the proof 1 of the theorem.

Proof 2: A second method of proving this theorem can also be done as follows. Let $m_1(.)$, $m_2(.)$, ..., $m_s(.)$, for $s \geq 3$, be bba's of the sources of information to combine and denote $\mathcal{F} = \{F_1, F_2, \dots, F_n\}$, for $n \geq 2$, the set of all focal elements. All sources give only binary masses, i.e. $m_k(F_l) = 0$ or $m_k(F_l) = 1$ for any $k \in \{1, 2, \dots, s\}$ and any $l \in \{1, 2, \dots, n\}$. Since each F_i, $1 \leq i \leq n$, is a focal element, there exists at least a bba $m_{i_o}(.)$ such that $m_{i_o}(F_i) = 1$, otherwise (i.e. if all sources gave the mass of F_i be equal to zero) F_i would not be focal. Without reducing the generality of the theorem, we can regroup the masses (since we combine all of them at once, so their order doesn't matter), as in Table II. Of course $i_1 + i_2 + \dots + i_n = s$, since the s bba's are the same but reordered, and $i_1 \geq 1$, $i_2 \geq 1$, ..., and $i_n \geq 1$. The total conflicting mass according to the theorem hypothesis $m_{1,2,\dots,s}(\emptyset)$ is 1. With the PCR6 fusion rule we transfer the conflict mass back to focal elements F_1, F_2, $\dots F_n$ respectively according to PCR

Table II. LIST OF REORDERED BINARY BBA'S.

bba's \ Focal elem.	F_1	F_2	...	F_n	$\emptyset$
$m_{r_1}(.)$	1	0	...	0	0
$m_{r_2}(.)$	1	0	...	0	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{r_{i_1}}(.)$	1	0	...	0	0
$m_{s_1}(.)$	0	1	...	0	0
$m_{s_2}(.)$	0	1	...	0	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{s_{i_2}}(.)$	0	1	...	0	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{u_1}(.)$	0	0	...	1	0
$m_{u_2}(.)$	0	0	...	1	0
⋮	⋮	⋮	⋮	⋮	⋮
$m_{u_{i_n}}(.)$	0	0	...	1	0
$m_{1,2,\dots,s}(.)$	0	0	...	0	1

principle such that:

$$\frac{x_{F_1}}{\underbrace{1+1+\dots+1}_{i_1 \text{ times}}} = \frac{x_{F_2}}{\underbrace{1+1+\dots+1}_{i_2 \text{ times}}} = \dots = \frac{x_{F_n}}{\underbrace{1+1+\dots+1}_{i_n \text{ times}}} = \frac{m_{1,2,\dots,s}(\emptyset)}{i_1+i_2+\dots+i_n} = \frac{1}{s}$$

whence $x_{F_1} = i_1/s$, $x_{F_2} = i_2/s$, ..., $x_{F_n} = i_n/s$. Therefore $m_{1,2,\dots,s}^{PCR6}(F_1) = i_1/s$, $m_{1,2,\dots,s}^{PCR6}(F_2) = i_2/s$, $\dots m_{1,2,\dots,s}^{PCR6}(F_n) = i_n/s$. But averaging the masses $m_1(.)$, $m_2(.)$, ..., $m_s(.)$ is equivalent to averaging each column of F_1, F_2, $\dots F_n$. Hence average of column F_1 is i_1/s, average of column F_2 is i_2/s, ..., average of column F_n is i_n/s. Therefore, in case of binary bba's which are globally totally conflicting, PCR6 rule is equal to the Averaging Rule. This completes the proof 2 of the theorem.

Note that using PCR5 fusion rule, we also transfer the total conflicting mass that is equal to 1 to F_1, F_2, ..., F_n respectively, but we replace the addition "+" with the multiplication "$\cdot$" in the above proportionalizations:

$$\frac{x_{F_1}}{\underbrace{1 \cdot 1 \cdot \dots \cdot 1}_{i_1 \text{ times}}} = \frac{x_{F_2}}{\underbrace{1 \cdot 1 \cdot \dots \cdot 1}_{i_2 \text{ times}}} = \dots = \frac{x_{F_n}}{\underbrace{1 \cdot 1 \cdot \dots \cdot 1}_{i_n \text{ times}}} = \frac{m_{1,2,\dots,s}(\emptyset)}{\underbrace{1+1+\dots+1}_{n \text{ times}}} = \frac{1}{n}$$

so that $x_{F_1} = 1/n$, $x_{F_2} = 1/n$, ..., $x_{F_n} = 1/n$ and therefore

$$m_{1,2,\dots,s}^{PCR5}(F_1) = m_{1,2,\dots,s}^{PCR5}(F_2) = \dots = m_{1,2,\dots,s}^{PCR5}(F_n) = 1/n$$

Corollary 1: When $s \geq 2$ sources of evidences provide binary bba's on G^Θ with at least two focal elements, and all focal elements are disjoint two by two, then PCR6 fusion rule coincides with the Averaging Rule.

This Corollary is true because if all focal elements are disjoint two by two then the total conflict is equal to 1.

Examples 2: where PCR6 rule equals the Averaging Rule.

Let's consider the frame $\Theta = \{A, B\}$ with Shafer's model and the bba's to combine as given in Table III.

Table III. LIST OF BBA'S TO COMBINE FOR EXAMPLE 2.

bba's \ Focal elem.	A	B	$A \cup B$	$A \cap B = \emptyset$
$m_1(.)$	1	0	0	
$m_2(.)$	0	1	0	
$m_3(.)$	0	0	1	
$m_{1,2,3}(.)$	0	0	0	1

Since we have binary masses, and their total conflict is 1, we expect getting the same result for PCR6 and the Averaging Rule according to our Theorem 1. The PCR principle gives us

$$\frac{x_A}{1} = \frac{y_B}{1} = \frac{z_{A\cup B}}{1} = \frac{m_{1,2,3}(\emptyset)}{1+1+1} = \frac{1}{3}$$

Hence $x_A = y_B = z_{A\cup B} = \frac{1}{3}$, so that

$$m_{1,2,3}^{PCR6}(A) = m_{1,2,3}(A) + x_A = 0 + \frac{1}{3} = \frac{1}{3}$$
$$m_{1,2,3}^{PCR6}(B) = m_{1,2,3}(B) + y_B = 0 + \frac{1}{3} = \frac{1}{3}$$
$$m_{1,2,3}^{PCR6}(A \cup B) = m_{1,2,3}(A \cup B) + z_{A\cup B} = 0 + \frac{1}{3} = \frac{1}{3}$$

Interestingly, PCR5 gives the same result as PCR6 in this case since one makes the same proportionalizations as for PCR6. Using the Averaging Rule (2), we get

$$m_{1,2,3}^{Average}(A) = \frac{1}{3} \cdot (1+0+0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(B) = \frac{1}{3} \cdot (0+1+0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(A \cup B) = \frac{1}{3} \cdot (0+0+1) = \frac{1}{3}$$

So we see that PCR6 rule equals the Averaging Rule as proved in the theorem because the bba's are binary and the intersection of all focal elements is empty since $A \cap B \cap (A \cup B) = \emptyset \cap (A \cup B) = \emptyset$ because $A \cap B = \emptyset$ since Shafer's model has been assumed for the frame Θ.

Examples 3: where PCR6 differs from the Averaging Rule.

Let's consider the frame $\Theta = \{A, B, C\}$ with Shafer's model and the bba's to combine as given in Table IV.

Table IV. LIST OF BBA'S TO COMBINE FOR EXAMPLE 3.

bba's \ Focal elem.	A	$A \cup B$	$A \cup B \cup C$	$\emptyset$
$m_1(.)$	1	0	0	
$m_2(.)$	0	1	0	
$m_3(.)$	0	0	1	
$m_{1,2,3}(.)$	1	0	0	

Clearly, in this case the focal elements are nested and the condition on emptiness of intersection of all focal elements is not satisfied because one has $A \cap (A \cup B) \cap (A \cup B \cup C) = A \neq \emptyset$, so that the theorem cannot be applied in such case. The total conflicting mass is not 1. One can verify in such example that PCR6 rule differs from the Averaging Rule because one gets

$$m_{1,2,3}^{PCR6}(A) = m_{1,2,3}(A) = 1$$
$$m_{1,2,3}^{PCR6}(A \cup B) = m_{1,2,3}(A \cup B) = 0$$
$$m_{1,2,3}^{PCR6}(A \cup B \cup C) = m_{1,2,3}(A \cup B \cup C) = 0$$

since there is no conflicting mass to redistribute to apply PCR principle, whereas the averaging fusion rule gives

$$m_{1,2,3}^{Average}(A) = \frac{1}{3} \cdot (1+0+0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(A \cup B) = \frac{1}{3} \cdot (0+1+0) = \frac{1}{3}$$
$$m_{1,2,3}^{Average}(A \cup B \cup C) = \frac{1}{3} \cdot (0+0+1) = \frac{1}{3}$$

Examples 4 (Bayesian non-binary bba's): where PCR6 differs from the Averaging Rule.

Let's consider the frame $\Theta = \{A, B\}$ with Shafer's model and the Bayesian bba's to combine as given in Table V.

Table V. LIST OF BBA'S TO COMBINE FOR EXAMPLE 4.

bba's \ Focal elem.	A	B	$A \cap B = \emptyset$
$m_1(.)$	0.2	0.8	0
$m_2(.)$	0.6	0.4	0
$m_3(.)$	0.7	0.3	0
$m_{1,2,3}(.)$	0.084	0.096	0.820

The total conflicting mass $m_{1,2,3}(A \cap B = \emptyset) = 0.82 = 1 - m_1(A)m_2(A)m_3(A) - m_1(B)m_2(B)m_3(B)$ equals the sum of partial conflicting masses that will be redistributed through PCR principle in PCR6

$$m_{1,2,3}(A \cap B = \emptyset) = \underbrace{m_1(A)m_2(B)m_3(B)}_{0.024} + \underbrace{m_2(A)m_1(B)m_3(B)}_{0.144} + \underbrace{m_3(A)m_1(B)m_2(B)}_{0.224} + \underbrace{m_1(B)m_2(A)m_3(A)}_{0.336} + \underbrace{m_2(B)m_1(A)m_3(A)}_{0.056} + \underbrace{m_3(B)m_1(A)m_2(A)}_{0.036} = 0.82$$

Applying PCR principle for each of these six partial conflicts, one gets:

- for $m_1(A)m_2(B)m_3(B) = 0.2 \cdot 0.4 \cdot 0.3 = 0.024$

  $$\frac{x_1(A)}{0.2} = \frac{y_1(B)}{0.4+0.3} = \frac{0.024}{0.2+0.3+0.4}$$

  whence $x_1(A) \approx 0.005333$ and $y_1(B) \approx 0.018667$.
- for $m_2(A)m_1(B)m_3(B) = 0.6 \cdot 0.8 \cdot 0.3 = 0.144$

  $$\frac{x_2(A)}{0.6} = \frac{y_2(B)}{0.8+0.3} = \frac{0.144}{0.6+0.8+0.3}$$

  whence $x_2(A) \approx 0.050824$ and $y_2(B) \approx 0.093176$.
- for $m_3(A)m_1(B)m_2(B) = 0.7 \cdot 0.8 \cdot 0.4 = 0.224$

  $$\frac{x_3(A)}{0.7} = \frac{y_3(B)}{0.8+0.4} = \frac{0.224}{0.7+0.8+0.4}$$

  whence $x_3(A) \approx 0.082526$ and $y_3(B) \approx 0.141474$.
- for $m_1(B)m_2(A)m_3(A) = 0.8 \cdot 0.6 \cdot 0.7 = 0.336$

  $$\frac{x_4(A)}{0.6+0.7} = \frac{y_4(B)}{0.8} = \frac{0.336}{0.8+0.6+0.7}$$

  whence $x_4(A) \approx 0.208000$ and $y_4(B) \approx 0.128000$.

- for $m_2(B)m_1(A)m_3(A) = 0.4 \cdot 0.2 \cdot 0.7 = 0.056$

$$\frac{x_5(A)}{0.2+0.7} = \frac{y_5(B)}{0.4} = \frac{0.056}{0.4+0.2+0.7}$$

whence $x_5(A) \approx 0.038769$ and $y_5(B) \approx 0.017231$.

- for $m_3(B)m_1(A)m_2(A) = 0.3 \cdot 0.2 \cdot 0.6 = 0.036$

$$\frac{x_6(A)}{0.2+0.6} = \frac{y_6(B)}{0.3} = \frac{0.036}{0.3+0.2+0.6}$$

whence $x_6(A) \approx 0.026182$ and $y_6(B) \approx 0.009818$.

Therefore, with PCR6 one finally gets

$$m_{1,2,3}^{PCR6}(A) = 0.084 + \sum_{i=1}^{6} x_i(A) = 0.495634$$
$$m_{1,2,3}^{PCR6}(B) = 0.096 + \sum_{i=1}^{6} y_i(A) = 0.504366$$

whereas the Averaging Rule (2) will give us

$$m_{1,2,3}^{Average}(A) = \frac{1}{3} \cdot (0.2+0.6+0.7) = \frac{1.5}{3} = 0.5$$
$$m_{1,2,3}^{Average}(B) = \frac{1}{3} \cdot (0.8+0.4+0.3) = \frac{1.5}{3} = 0.5$$

In this example, the intersection of focal elements is empty but the bba's to combine are not binary. Therefore the total conflict between sources is not total and the theorem doesn't apply and so PCR6 results differ from the Averaging Rule.

It however can happen that in some very particular symmetric cases PCR6 coincides with the Averaging Rule. For example, if we consider the bba's as given in the Table VI. In such case the opinion of source #1 totally balances opinion of source #3, and the opinion of source #2 cannot support A more than B (and reciprocally), so that the fusion problem is totally symmetrical. In this example, it is expected that the final fusion result should commit an equal mass of belief to A and to B. And indeed, it can be easily verified that one gets in such case

$$m_{1,2,3}^{PCR6}(A) = m_{1,2,3}^{Average}(A) = 0.5$$
$$m_{1,2,3}^{PCR6}(B) = m_{1,2,3}^{Average}(B) = 0.5$$

which makes perfectly sense. Note that the Averaging Rule provides same result on example 4 which is somehow questionable because example 4 doesn't present an inherent symmetrical structure. In our opinion PCR6 presents the advantage to respond more adequately to the change of inherent internal structure (asymmetry) of bba's to combine, which is not well captured by the simple averaging fusion rule.

Table VI. A BAYESIAN NON-BINARY SYMMETRIC EXAMPLE.

bba's \ Focal elem.	A	B	$A \cap B = \emptyset$
$m_1(.)$	0.2	0.8	0
$m_2(.)$	0.5	0.5	0
$m_3(.)$	0.8	0.2	0
$m_{1,2,3}(.)$	0.08	0.08	0.84

IV. APPLICATION TO PROBABILITY ESTIMATION

Let's review a simple coin tossing random experiment. When we flip a coin [13], there are two possible outcomes. The coin could land showing a head (H) or a tail (T). The list of all possible outcomes is called the sample space and correspond to the frame $\Theta = \{H, T\}$. There exist many interpretations of probability [14] that are out of the scope of this paper. We focus here on the estimation of the probability measure $P(H)$ of a given coin (biased or not) based on n outcomes of a coin tossing experiment. The long-run frequentist interpretation of probability [15] considers that the probability of an event A is its relative frequency of occurrence over time after repeating the experiment a large number of times under similar circumstances, that is

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n} \tag{12}$$

where $n(A)$ denotes the number of occurrences of an event A in $n > 0$ trials. In practice however, we usually estimate the probability of an event A based only on a limited number of data (observations) that are available, and so we estimate the idealistic $P(A)$ defined in (12), by classical Laplace's probability definition

$$\hat{P}(A|n(A), n) = \frac{n(A)}{n} \tag{13}$$

Naturally, $\hat{P}(A) \geq 0$ because $n(A) \geq 0$ and $n > 0$, and $\hat{P}(A) \leq 1$ because we cannot get $n(A) > n$ in a series of n trials. $P(A) + P(\bar{A}) = 1$ because $\frac{n(A)}{n} + \frac{n(\bar{A})}{n} = \frac{n(A)}{n} + \frac{n-n(A)}{n} = 1$ where $\bar{A}$ is the complement of A in the sample space.

It is interesting to note that the classical estimation of the probability measure given by (13) corresponds in fact to the simple averaging fusion rule of distinct pieces of evidence represented by binary masses. For example, let's take a coin and flip it $n = 8$ times and assume for instance that we observe the following series of outcomes $\{o_1 = H, o_2 = H, o_3 = T, o_4 = H, o_5 = T, o_6 = H, o_7 = H, o_8 = T\}$, so that $n(H) = 5$ and $n(T) = 3$. Then these observations can be associated with distinct sources of evidences providing to the following basic (binary) belief assignments:

Table VII. OUTCOMES OF A COIN TOSSING EXPERIMENT.

bba's \ Focal elem.	H	T
$m_1(.)$	1	0
$m_2(.)$	1	0
$m_3(.)$	0	1
$m_4(.)$	1	0
$m_5(.)$	0	1
$m_6(.)$	1	0
$m_7(.)$	1	0
$m_8(.)$	0	1

It is clear that the probability estimate in (13) equals the averaging fusion rule (2) and in such example because

$$\begin{aligned}\hat{P}(H|\{o_1, o_2, \ldots, o_8\}) &= \frac{n(H)}{n} = \frac{5}{8} \qquad \text{by eq. (13)} \\ &= \frac{1}{8}(1+1+0+1+0+1+1+0) \\ &= m_{1,2,\ldots,8}^{Average}(H) \qquad \text{by eq. (2)}\end{aligned}$$

$$\begin{aligned}\hat{P}(T|\{o_1,o_2,\ldots,o_8\}) &= \frac{n(T)}{n} = \frac{3}{8} \qquad \text{by eq. (13)}\\ &= \frac{1}{8}(0+0+1+0+1+0+0+1)\\ &= m_{1,2,\ldots,8}^{Average}(T) \qquad \text{by eq. (2)}\end{aligned}$$

Because all the bba's to combine here are binary and are in total conflict, our theorem 1 of Section III applies, and PCR6 fusion rule in this case coincides with the averaging fusion rule. Therefore we can use PCR6 to estimate the probabilities that the coin will land on H or T at the next toss given the series of observations. More precisely,

$$\begin{cases} m_{1,2,\ldots,8}^{PCR6}(H) = m_{1,2,\ldots,8}^{Average}(H) = \hat{P}(H|\{o_1,o_2,\ldots,o_8\}) \\ m_{1,2,\ldots,8}^{PCR6}(T) = m_{1,2,\ldots,8}^{Average}(T) = \hat{P}(T|\{o_1,o_2,\ldots,o_8\}) \end{cases}$$

We must insist on the fact that Dempster-Shafer (DS) rule of combination (4) cannot be used at all in such very simple case to estimate correctly the probability measure because DS rule doesn't work (because of division by zero) in total conflicting situations. PCR5 rule can be applied to combine these 8 bba's but is unable to provide a consistent result with the classical probability estimates because one will get

$$\frac{x_H}{1\cdot 1\cdot 1\cdot 1\cdot 1} = \frac{y_T}{1\cdot 1\cdot 1} = \frac{m_{1,2,\ldots,8}(\emptyset)}{(1\cdot 1\cdot 1\cdot 1\cdot 1)+(1\cdot 1\cdot 1)} = \frac{1}{1+1} = 0.5$$

and therefore the PCR5 fusion result is

$$\begin{cases} m_{1,2,\ldots,8}^{PCR5}(H) = x_H = 0.5 \neq (m_{1,2,\ldots,8}^{PCR6}(H) = 5/8) \\ m_{1,2,\ldots,8}^{PCR5}(T) = y_T = 0.5 \neq (m_{1,2,\ldots,8}^{PCR6}(T) = 3/8) \end{cases}$$

Remark: The PCR6 fusion result is valid if and only if PCR6 rule is applied globally, and not sequentially. If PCR6 is sequentially applied, it becomes equivalent with PCR5 sequentially applied and it will generate incorrect results for combining $s > 2$ sources of evidence. Because of the ability of PCR6 to estimate frequentist probabilities in a random experiment, we strongly recommend PCR6 rather than PCR5 as soon as $s \geq 2$ bba's have to be combined altogether.

V. Conclusions and challenge

In this paper, we have proved that PCR6 fusion rule coincides with the Averaging Rule when the bba's to combine are binary and in total conflict. Because of such nice property, PCR6 is able to provide a frequentist probability measure of any event occurring in a random experiment, contrariwise to other fusion rules like DS rule, PCR5 rule, etc. Except the Averaging Rule of course since it is the basis of the frequentist probability interpretation. In a more general context with non-binary bba's, PCR6 is quite complicate to apply to combine globally $s > 2$ sources of evidences, and a general recursive formula of PCR6 would be very convenient. It can be mathematically reformulated as follows: Let R be a fusion rule and assume we have s sources that provide m_1, m_2, ..., m_{s-1}, m_s respectively on a fusion space G^Θ. Find a function (or an operator) T such that: $T(R(m_1,m_2,\ldots m_{s-1}),m_s) = R(m_1,m_2,\ldots,m_{s-1},m_s)$, or by simplifying the notations $T(R_{s-1},m_s) = R_s$, where R_i means the fusion rule R applied to i masses all together. For example, if R equals the Averaging Rule, the function T is defined according to the relation (3) by $T(R_{s-1},m_s) = \frac{s-1}{s}R_{s-1} + \frac{1}{s}m_s = R_s$, and if R equals DS rule one has $T(R_{s-1},m_s) = DS(R_{s-1},m_s)$ because of the associativity of DS rule. What is the T operator associated with PCR6? Such very important open challenging question is left for future research works.

References

[1] G. Shafer, A mathematical theory of evidence, Princeton Univ. Press, Princeton, NJ, U.S.A., 1976.

[2] F. Smarandache, J. Dezert (Editors), Advances and applications of DSmT for information fusion, American Research Press, Rehoboth, NM, U.S.A., Vol. 1–3, 2004–2009. http://fs.gallup.unm.edu//DSmT.htm

[3] X.R. Li, Probability, random signals and statistics, CRC Press, 1999.

[4] P.E. Pfeiffer, Applied probability, Connexions Web site. http://cnx.org/content/col10708/1.6/

[5] L.A. Zadeh, On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, CA, U.S.A., 1979.

[6] L.A. Zadeh, Book review: A mathematical theory of evidence, The Al Magazine, Vol. 5, No. 3, pp. 81-83, 1984.

[7] L.A. Zadeh, A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, The Al Magazine, Vol. 7, No. 2, 1986.

[8] J. Dezert, P. Wang, A. Tchamova, On the validity of Dempster-Shafer theory, Proc. of Fusion 2012 Int. Conf., Singapore, July 9–12, 2012.

[9] A. Tchamova, J. Dezert, On the behavior of Dempster's Rule of combination and the foundations of Dempster-Shafer theory, (Best paper awards), 6th IEEE Int. Conf. on Int. Syst. (IS '12), Sofia, Bulgaria, Sept. 6–8, 2012.

[10] F. Smarandache, J. Dezert, J.-M. Tacnet, Fusion of sources of evidence with different importances and reliabilities, Proc. of Fusion 2010 Int. Conf., Edinburgh, UK, July 26–29, 2010.

[11] http://bfas.iutlan.univ-rennes1.fr/wiki/index.php/Toolboxs

[12] F. Smarandache, J. Dezert, Importance of sources using the repeated fusion method and the proportional conflict redistribution rules #5 and #6, in Multispace & Multistructure. Neutrosophic transdisciplinarity (100 collected papers of sciences), Vol. IV, pp. 349–354, North-European Scientific Publishers Hanko, Finland, 2010.

[13] http://www.random.org/coins/ or http://shazam.econ.ubc.ca/flip/

[14] Free online Stanford Encyclopedia of Philosophy. http://plato.stanford.edu/entries/probability-interpret/

[15] R. Von Mises, Probability, statistics and the truth, Dover, New York, Second revised English Edition, 1957.

Examples where the Conjunctive and Dempster's Rules are Insensitive

Florentin Smarandache
Jean Dezert
Valeri Kroumov

Originally published as *Smarandache F., Kroumov V., Dezert J., Examples where the conjunctive and Dempster's rules are insensitive,* Proc. of 2013 International Conference on Advanced Mechatronic Systems, Luoyang, China, Sept. 25-27, 2013, and reprinted with permission.

***Abstract*—In this paper we present several counter-examples to the Conjunctive rule and to Dempster rule of combinations in information fusion.**

Keywords— conjunctive rule, Dempster rule, DSmT, counter-examples to Conjunctive rule, counter-examples to Dempster rule, information fusion

I. INTRODUCTION

In *Counter-Examples to Dempster's Rule of Combination* {Ch. 5 of *Advances and Applications to DSmT on Information Fusion,* Vol. I, pp. 105-121, 2004} [1], J. Dezert, F. Smarandache, and M. Khoshnevisan have presented several classes of fusion problems which could not be directly approached by the classical mathematical theory of evidence, also known as Dempster-Shafer Theory (DST), either because Shafer's model for the frame of discernment was impossible to obtain, or just because Dempster's rule of combination failed to provide coherent results (or no result at all). We have showed and discussed the potentiality of the DSmT combined with its classical (or hybrid) rule of combination to attack these infinite classes of fusion problems.
We have given general and concrete counter-examples for Bayesian and non-Bayesian cases.

In this article we construct new classes where both the conjunctive and Dempster's rule are insensitive.

II. DEZERT-TCHAMOVA COUNTER-EXAMPLE

In [2], J. Dezert and A. Tchamova have introduced for the first time the following counter-example with some generalizations. This first type of example has then been discussed in details in [3,4] to question the validity of foundations of Dempster-Shafer Theory (DST). In the next sections of this short paper, we provide more counter-examples extending this idea. Let the frame of discernment $\Theta = \{A, B, C\}$, under Shafer's model (i.e. all intersections are empty), and $m_1(.)$ and $m_2(.)$ be two independent sources of information that give the below masses:

Focal Elements	*A*	*C*	$A \cup B$	$A \cup B \cup C$
m_1	a	0	$1-a$	0
m_2	0	$1-b_1-b_2$	b_1	b_2

Table 1

where the parameters $a, b_1, b_2 \in [0,1]$, and $b_1+b_2 \leq 1$.

Applying the conjunctive rule, in order to combine $m_1 \oplus m_2 = m_{12}$, one gets:

$$m_{12}(A) = a(b_1+b_2) \quad (1)$$

$$m_{12}(C) = 0 \quad (2)$$

$$m_{12}(A \cup B) = (1-a)(b_1+b_2) \quad (3)$$

$$m_{12}(A \cup B \cup C) = 0 \quad (4)$$

and the conflicting mass

$$m_{12}(\phi) = 1-b_1-b_2 = K_{12}. \quad (5)$$

After normalizing by diving by $1-K_{12} = b_1+b_2$ one gets Demspter's rule result $m_{DS}(.)$:

$$m_{DS}(A) = \frac{m_{12}(A)}{1 \quad K_{12}} = \frac{a(b_1+b_2)}{b_1+b_2} = a = m_1(A)$$

$$m_{DS}(A \cup B) = \frac{m_{12}(A \cup B)}{1 \quad K_{12}} = \frac{(1-a)(b_1+b_2)}{b_1+b_2} = 1-a = m_1(A \cup B) \quad (6)$$

Counter-intuitively after combining two sources of information, $m_1(.)$ and $m_2(.)$, with Dempster's rule, the result does not depend at all on $m_2(.)$. Therefore Dempster's rule is insensitive to $m_2(.)$ no matter what the parameters a, b_1, b_2 are equal to.

III. FUSION SPACE

In order to generalize this counter-example, let's start by defining the fusion space.

Let Θ be a frame of discernment formed by *n* singletons Ai, defined as:

$$\Theta = \{\phi_1, \phi_2, ..., \phi_n\}, n \geq 2, \quad (7)$$

and its Super-Power Set (or fusion space):

$$S^\Theta = (\Theta, \cup, \cap, C) \quad (8)$$

which means the set Θ closed under union $\cup$, intersection $\cap$, and respectively complement C.

IV. ANOTHER CLASS OF COUNTER-EXAMPLES TO DEMPSTER'S RULE

Let $A_1, A_2, ..., A_p \in S^\Theta \setminus \{I_t, \phi\}$, for $p \geq 1$, such that $Ai \cap Aj = \phi$ for $i \neq j$, where I_t is the total ignorance $(A_1 \cup A_2 \cup ... \cup A_n)$, and ϕ is the empty set.

Therefore each A_i, for $i \in \{1, 2, ..., p\}$, can be either a singleton, or a partial ignorance (union of singletons), or an intersection of singletons, or any element from the Super-Power Set S^Θ (except the total ignorance or the empty set), i.e. a general element in the set theory that is formed by the operators $\cup, \cap, C$.

Let's consider two sources $m_1(.)$ and $m_2(.)$ defined on S^Θ:

	A_1	A_2	...	A_p	I_t
m_1	a_1	a_2	...	a_p	0
m_2	b	b	...	b	$1-p\cdot b$

where of course all $a_i \in [0, 1]$ and $a_1 + a_2 + ... + a_p = 1$,
also b and $1-p\cdot b \in [0, 1]$.

$m_1(.)$ can be Bayesian or non-Bayesian depending on the way we choose the focal elements $A_1, A_2, ..., A_p$.

We can make sure $m_2(.)$ is not the uniform basic believe assignment by setting $b \neq 1-p\cdot b$.

Let's use the conjunctive rule for $m_1(.)$ and $m_2(.)$:

$m_{12}(A_i) = m_1(A_i)m_2(A_i) + [m_1(A_i)m_2(I_t) + m_1(I_t)m_2(A_i)] = a_i\cdot b + [a_i\cdot(1-p\cdot b) + 0\cdot b] = a_i\cdot(1-p\cdot b+b)$, for all $i \in \{1, 2, ..., p\}$. (9)

It is interesting to finding out, according to the Conjunctive Rule, that the conflict of the above two sources does not depend on $m_1(.)$ at all, but only on $m_2(.)$, which is abnormal:

$$K_{12} = \sum_{i=1}^{p}\sum_{\substack{j=1 \\ j\neq i}}^{p} m_1(A_i)m_2(A_j) = \sum_{i=1}^{p}\sum_{\substack{j=1 \\ j\neq i}}^{p} a_i\cdot b = \sum_{i=1}^{p}(p-1)a_i\cdot b = (p-1)b\sum_{i=1}^{p} a_i = (p-1)b. \quad (10)$$

Therefore even the feasibility of the Conjunctive Rule is questioned.

When we normalize, as in Dempster's Rule, by dividing all $m_{12}(.)$ masses by the common factor $1-K = 1-p\cdot b+b$, we actually get: $m_1 \oplus m_2 = m_1$! So, $m_2(.)$ makes no impact on the fusion result according to Dempster's Rule, which is not normal.

V. MORE GENERAL CLASS OF COUNTER-EXAMPLES TO DEMPSTER'S RULE

Let's consider $r+1$ sources: the previous $m_1(.)$ and respectively various versions of the previous $m_2(.)$:

	A_1	A_2	...	A_p	I_t
m_1	a_1	a_2	...	a_p	0
m_{21}	b_1	b_1	...	$b1$	$1-p\cdot b_1$
m_{22}	b_2	b_2	...	b_2	$1-p\cdot b_2$
.					
.					
.					
m_{2r}	b_r	b_r	...	b_r	$1-p\cdot b_r$

where of course all $a_i \in [0, 1]$ and $a_1 + a_2 + ... + a_p = 1$,
also all b_j and $1-p\cdot b_j \in [0, 1]$, for $j \in \{1, 2, ..., r\}$. (11)

Now, if we combine $m_1 \oplus m_{21} \oplus m_{22} \oplus ... \oplus m_{2r} = m_1$. Therefore all r sources $m_{21}(.), m_{22}(.), ..., m_{2r}(.)$ have no impact on the fusion result!

Interesting particular examples can be found in this case.

VI. SHORT GENERALIZATION OF DEZERT-TCHAMOVA COUNTER-EXAMPLE

Let's consider four focal elements A, B_1, B_2, B_3, such that $A \cap B_i = \phi$ for $i \in \{1,2,3\}$, and B_1, B_2, B_3 are nested, i.e. $B_1 \subset B_2 \subset B_3$, and two masses, where of course $b_1+b_2 = 1$ and $c_1+c_2+c_3 = 1$, and all $b_1, b_2, c_1, c_2, c_3 \in [0, 1]$:

	A	B_1	B_2	B_3
m_1	0	b_1	b_2	0
m_2	c_1	0	c_2	c_3
---	---	---	---	---
m_{12}	0	$b_1(1-c_1)$	$b_2(1-c_1)$	0

and the conflict $K_{12} = c_1(b_1+b_2)=c_1$

m_D	0	b_1	b_2	0

a) This generalization permits the usefulness of hybrid models, for example one may have the frame of discernment of exclusive elements {A, B, C}, where $B_1 = B \cap C$, $B_2 = B$, and $B_3 = B \cup C$.
b) Other interesting particular cases may be derived from this short generalization.

VII. PARTICULAR COUNTER-EXAMPLE TO THE CONJUNCTIVE RULE AND DEMPSTER'S RULE

For example let $\Theta = \{A, B, C\}$, in Shafer's model. We show that the conflicts between sources are not correctly reflected by the conjunctive rule, and that a certain non-vacuous non-uniform source is ignored by Dempster's rule.

Let's consider the masses:

	A	B	C	$A \cup B \cup C$	
m_1	1	0	0	0	*(the most specific mass)*
m_2	1/3	1/3	1/3	0	*(very unspecific mass)*
m_3	0.6	0.4	0	0	*(mass between the very unspecific and the most specific masses)*
m_0	0.2	0.2	0.2	0.4	*(not vacuous mass, not uniform mass)*

Then the conflict $K_{10} = 0.4$ between $m_1(.)$ and $m_0(.)$ is the same as the conflict K_{20} between $m_2(.)$ and $m_0(.)$, and similarly the same as the conflict K_{30} between $m_3(.)$ and $m_0(.)$, which is not normal, since $m_1(.)$ is the most specific mass while $m_2(.)$ is the most unspecific mass.

Let's check other thing combining two sources using Dempster's rule:
$m_1 \oplus m_o = m_1$, $m_2 \oplus m_o = m_2$, $m_3 \oplus m_0 = m_3$, which is not normal.

In order to get the "normal behavior" we combine $m_1(.)$ and $m_0(.)$ with *PCR5*, and similarly for others: $m_2(.)$ combined with $m_0(.)$, and $m_3(.)$ combined with $m_o(.)$.

In order to know what should have been the "normal behavior" for the conflict (the initial conflict was $K_{10} = 0.4$), let's make a small change to $m_0(.)$ as below:

	A	B	C	$A \cup B \cup C$	
m_1	1	0	0	0	*(the most specific mass)*
m_2	1/3	1/3	1/3	0	*(very unspecific mass)*
m_3	0.6	0.4	0	0	*(mass between the very unspecific and the most specific masses)*
m_0	0.3	0.2	0.1	0.4	*(not vacuous mass, not uniform mass)*

$K_{10} = 0.30$
$K_{20} = 0.40$
$K_{30} = 0.34$

Now, the conflicts are different.

VIII. CONCLUSION

We showed in this paper that: first the conflict was the same, no matter what was one of the sources (and it is abnormal that a non-vacuous non-uniform source has no impact on the conflict), and second that the result using Dempster's rule is not all affected by a non-vacuous non-uniform source of information.

Normally, the most specific mass (*bba*) should dominate the fusion result.

Therefore, the conflicts between sources are not correctly reflected by the conjunctive rule, and certain non-vacuous non-uniform sources are ignored by Dempster's rule in the fusion process.

REFERENCES

[1] F. Smarandache, J. Dezert, "Advances and Applications of DSmT for Information Fusion," Vols. I, II, III, American Res. Press, Rehoboth, 2004, 2006, respectively 2009; http://fs.gallup.unm.edu/DSmT.htm.

[2] J. Dezert, A. Tchamova, "On the behavior of Dempster's rule of combination," Presented at the spring school on Belief Functions Theory and Applications (BFTA), Autrans, France, 4–8 April 2011 (http://hal.archives-ouvertes.fr/hal-00577983/).

[3] J. Dezert., P. Wang, A. Tchamova, "On The Validity of Dempster-Shafer Theory." Proceedings of the International Conference of Information Fusion, Singapore, July, 2012.

[4] A. Tchamova., J. Dezert, "On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory," (Best paper award), in Proc. of IEEE IS'2012, Sofia, Bulgaria, Sept. 6-8, 2012.

URREF Self-Confidence in Information Fusion Trust

Erik Blasch
Audun Jøsang
Jean Dezert
Paulo C. G. Costa
Anne-Laure Jousselme

Originally published as Blasch E., Josang A., Dezert J., Costa P., Jousselme A.-L., *URREF Self-Confidence in Information Fusion Trust,* in Proc. of Fusion 2014 Int Conf onInformation Fusion, Salamanca, Spain, July 7-10, 2014, and reprinted with permission.

Abstract *— The Uncertainty Representation and Reasoning Evaluation Framework (URREF) includes an ontology that represents concepts and criteria needed to evaluate the uncertainty management aspects of a fusion system. The URREF ontology defines self-confidence as a measure of the information credibility as evaluated by the sensor itself. The concept of confidence, which is not explicitly defined in the ontology at URREF, has been extensively explored in the literature about evaluation in information fusion systems (IFS). In this paper, we provide a discussion on confidence as it relates to the evaluation of IFS, and compare it with the existing concepts in the URREF ontology. Our goal is two-fold, since we address both the distinctions between confidence and self-confidence, as well as the implications of these differences when evaluating the impact of uncertainty to the decision-making processes supported byt the IFS. We illustrate the discussion with an example of decision making that involves signal detection theory, confusion matrix fusion, subjective logic, and proportional conflict redistribution. We argue that uncertainty can be minimized through confidence (information evidence) and self-confidence (source agent) processing, The results here seek to enrich the ongoing discussion at the ISIF's Evaluation of Techniques for Uncertainty Representation Working Group (ETURWG) on self-confidence and trust in information fusion systems design.*

Keywords: *Self-Confidence, Confidence, Trust, Level 5 Fusion, High-Level Information Fusion , PCR5/6, Subjective Logic*

I. INTRODUCTION

Information fusion aims to achieve uncertainty reduction through combining information from multiple complementary sources. The International Society of Information Fusion (ISIF) *Evaluation of Techniques of Uncertainty Reasoning Working Group* (ETURWG) was chartered to address the problem of evaluating fusion systems' approaches to representing and reasoning with uncertainty. The working group developed the *Uncertainty Representation and Reasoning Evaluation Framework (URREF)* [1]. Discussions during 2013 explored the notions of credibility and reliability [2]. One recent issue is the difference between confidence and self-confidence as related to the data, source, and processing. While agreement is not complete among the ETURWG, this paper seeks to provide one possible approach to relate the mathematical, semantic, and theoretical challenges of confidence analysis.

The key position of the paper is to analyze the practical differences in evaluating the two concepts. More specifically, self-confidence is mostly relevant to HUMINT, which makes its evaluation a primarily subjective; whereas confidence can be easily traced to machine data analysis, allowing for the use of objective metrics in its evaluation. That is, a machine can process large amounts of data to represent the state of the world, and the evaluation of how well uncertainty is captured in these processes can be traced to various objective metrics. In contrast, for a human to assess its own confidence on the credibility of his "data collection process" (i.e. self-confidence), he or she has to make a judgment on limited choices. Objective assessment is determined from the credibility of the reports, processing, and decisions. Typical approaches include artificial intelligence (AI) methods (e.g., Neural Networks), pattern recognition (e.g., Bayesian, wavelets), and automatic target exploitation (i.e., over sensor, target, and environment operating conditions [3]). Subjective analysis is a report quality opinion that factors in analysis (e.g., completeness; accuracy, and veracity), knowledge (e.g., representation, uncertainty, and reasoning), and judgment (e.g., intuition, experience, decision making) [4]. In terms of IFS support for decision-making, numerous methods have been explored, mainly from Bayesian reasoning, Dempster-Shafer Theory [5], Subjective Logic [6], DSmT [7], fuzzy logic and possibility theory; although it also includes research on approximating belief functions to subjective probability measures (BetP [8], DSmP [9]).

Figure 1 provides a framework for our discussion. The world contains some truth *T*, of which data is provided from different sources (*A*, *B*). Source *A* analysis goes to a machine agent for information fusion processing while source B goes to a human agent. Either the machine or the human can generate beliefs about the state of the world (either using qualitative or quantitative semantics). The combination of *A* and *B* is a subject of Level 5 Fusion (user refinement) [10, 11].

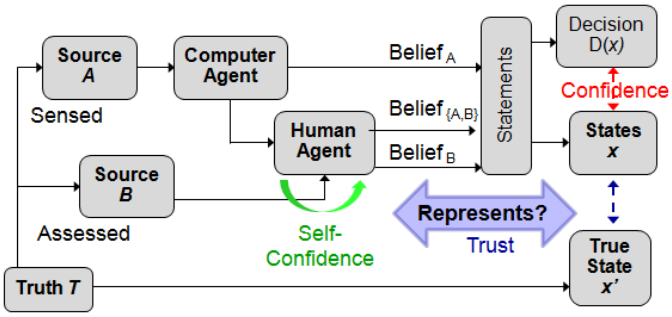


Figure 1 – Information Fusion System assessment of confidence (machine) and self-confidence (sensor or human agent).

On one hand, confidence is typically related to machine processing such as signal detection theory where self-confidence is associated with sensors (and humans) assessing their own capability. On the other hand, the manipulation of the data requires understanding of the source, and self-confidence is applicable for the cases in which the user can provide a self-assessment on how confident he is on its data. It is important to emphasize that are not dealing (at least not directly) with information veracity: even if the sensor (e.g., a human reporting on an event or providing assessment on a situation) considers the information as possible, and he trusts it, it could be false at the end (e.g., even in summer time we can have a cloudy day). That is, self-confidence assesses how much the author trust the information, but not necessarily that this information is false or true [1]. For this paper, we take the URREF ontology definition of self-confidence as implied in Figure 1. The rationale for this choice is that self-confidence and uncertainty are typically associated with humans whereas confidence has been typically used in signal detection. Fusion of beliefs ultimately relates to states of the world with a reported confidence that can be compared to a truth state. Debating on the overlaps in terminology would be welcomed to clarify these positions for the ETURWG and the information community as a whole.

From Figure 1, we note the importance of confidence as related from a decision to the estimated states. Self-confidence is within the human agent assessing their understanding (e.g. experience) that can also be combined with the computer agent. The issue at hand for a user is whether or not the machine analysis (or their own) state decision represents reality. The notion of reality comes from the fact that currently, there are many technical products that perceive the world for the user (e.g., video) from which the user must map a mental model to a physical model of what is being represented. Some cases are easy such as video of cars moving on a road [12]; however, others are complex such as cyber networks [13]. The example used through the rest of the paper requires High-Level Information Fusion (HLIF) of target detection from a machine and human [14, 15].

In designing computer-aided detection machines, it is desirable to provide intelligence amplification (IA) [16] where *Qualia* motivates subjective analysis as a relation between the human consciousness/self-awareness to external stimuli. Qualia is the internal perception of the subjective aspect of the human's perception of the stimuli. Knowing oneself can then be utilized to understand/evaluate the use of meaningful and relevant data in decision-making. The more that a sensor understands its Qualia [17], the better it will be in providing an assessment of its self-confidence in a report or on a decision. Qualia then encompasses an important component to uncertainty reasoning associated with subjective beliefs, trust, and self-confidence in decision making as a sense of intuition. Not surprisingly, these are natural discussion topics in Level 5 fusion ('user refinement'), which includes operator-machine collaboration [18], situation awareness/assessment displays [19], and trust [20]. In order to explore self-confidence on these issues, we need to look at the psychology literature on trust as it relates to self-confidence.

From data available on the web (e.g., twitter, documents), intelligent users need the capability to rapidly monitor and analyze event information over massive amounts of unstructured textual data [21]. Text from human sources is subjected to opinions, beliefs, and misperceptions, generating various forms of self-assumed self-confidence. In contrast, computer sensed data can be stochastic or deterministic, from which we have to coordinate the agent information. For example, with Gaussian observations generates stochastic probability analyses (e.g., Kalman Filter). However, structural information in the sensor models and sensitivities for a given state condition (which come from a deterministic ontology) could be used to improve the estimate [22]. This combination of both stochastic and deterministic decisions with uncertainty elements is usual in modeling and system deployment, and understanding its key aspects is a fertile area for producing better decision support from IFS.

A related example from the analysis of uncertainty is evidence assessment from *opinion makers*. Dempster-Shafer theory has been used in connection with Bayesian analysis for decision making [5].
Likewise, Jøsang [23] demonstrated how subjective analysis within Dempster-Shafer theory could be used to determine the weight of opinions. Ontologies such as the one used in the URREF must be able to account for the uncertainty of data and to model it qualitatively, semantically, and quantitatively [24]. Metrics such as quality of service (QoS) and quality or information (IQ) are example of tools that can support and enhance a modeling capability between ontologies and uncertainty analysis [25]. The rest of this paper includes Sect. II as an overview of self-confidence. Sect. III discusses the mathematical analysis. Sect. IV highlights subjective logic for opinion making. Sect. V is an example and Sect. VI provides conclusions.

II. URREF NOTIONS OF SELF-CONFIDENCE

The ETURWG has explored many topics as related to a systems analysis of information fusion, which includes characteristics of uncertainty with many unknowns [26, 27]. In this paper, we categorize the characteristics of uncertainty into four areas, shown in Table 1. Assuming that the flow of information first goes from an agent to evidence beliefs, and subsequently to fusion with knowledge representation, then these areas help understand the terminology. Note that the defined information fusion quality of service (QoS) parameters are in blue {timeliness, accuracy, throughput, and confidence}. These could also be measures of performance [28]. For measures of effectiveness [25], one needs to understand system robustness (e.g., consistency, completeness, correctness, integrity). Here we focus on the red terms as related to self-confidence and confidence.

Knowledge representation in IFS [29, 30] includes applying decision-making semantics to support the structuring of extracted information. One example is the use of well defined concepts (e.g. confirmed, probable, possible, doubtful, and improbable) to support information extraction with natural language processing (NLP) algorithms. As related to confidence and self-confidence, there is the notion of integrity.

Table 1: Characteristics of Uncertainty [3]

Agent	**Evidence**	**Algorithm**	**Representation**
Source	***Information***	***Fusion***	***Knowledge Reasoning***
		Scalability	Knowledge Handling
Objectivity	Relevance	Computational Cost	Simplicity
Observational Sensitivity	Conclusiveness	Adaptability	Expressiveness
Veracity (truthfulness)	Veracity (truth)	Traceability (pedigree)	Polarity
Secure	Ambiguity	Stability	Modality
Resilient			Genericity
Trust	Precision	**Throughput**	Tense
	Accuracy	**Timeliness**	
Reliability	**Credibility**	**Correctness**	**Completeness**
Self-Confidence	**Confidence**	**Consistency**	**Integrity**

[3] This table is presented to the ETURWG in this paper to support ongoing discussions on the categorization of types of uncertainty

Integrity for human agents is associated with their subjective accountability and consistency in making judgments. Integrity for a machine could be objective in the faithful representation and validity on the data [31].

Algorithm performance focus on the information fusion method. URREF criteria for evaluating it relates to how the uncertainty model performs operations with information. An example of related metrics is to assess uncertainty reduction by weighting good data over bad given conflicting data.

Evidence: From [2], we explored the weight of evidence (WOE) as a function of reliability, credibility, relevance, and completeness. In URREF, WOE assesses how well an uncertainty representation technique captures the impact of an input affecting the processing and output of the IFS.

Source: Self-confidence, while yet to have a clear definition in the engineering literature, is typically associated with trust.

A. Trust

Closely associated with subjective analysis is *trust* [32]. Trust includes many attributes for man-machine systems such as dependability (machine), competence (user), and application [33]. Trust is then related to machine processing (confidence) and human assessment (self-confidence). Trust in automation is a key attribute associated with machine-driven solutions. Human trust in automation determines a user's reliance on automation. In [32], they explored self-confidence defined as the user anticipatory (or post) performance with machines which impacts with trust in policy application.

Measuring trust as related to uncertainty is an open topic [34]. As a focus of discussion, we have a machine agent and a human agent of which a measure of trust comes from the uncertainty associated between the man-machine interactions. *Reliability trust* could be between human agents of which subjective probability is useful [35]. *Decision trust* could be between human agents or between a human and a machine and takes into account the risk associated with situation-dependent attitudes, attention, and workload of a human agent. The distinction between reliability and decision trust is important as related to self-confidence and confidence. This can be seen in Table 2, which depicts the main aspects for each of the six potential interactions between sensors..

Table 2: Trust Aspects in Sensor Interactions

	Human	**Others**	**Machine**
Human	Self-confidence	Reliability	Trustworthy
Machine	Trust	Credibility	Confidence

- *Human*: Individuals must provide introspection on their own analysis and interaction with a machine. Here we distinguish between self-confidence and trust. In this case, human agents must have self-confidence in themselves as well as trust in the machine.
- *Others*: With the explosion of the Internet, recent work has explored the uncertainty of human sensing, such as Twitter reports in social networks, showing humans as less calibrated and reliable in their sensing. Wang et al. [36, 37] developed an estimation approach for truth discovery in this domain. Another recent example explored the decision-making trust between humans interfacing through a machine. The user interface was shown to have a strong impact on trust, cooperation, and situation awareness [38]. As an interesting result, credibility resulted as the computer interaction afforded complete and incomplete information towards understanding both the machine and the user analysis.
- *Machine*: A large body of literature is devoted to network trust. Examples include the hardware, cyber networks [39], protocols and policies. Given the large amount of cyber attacks written by hackers, it comes down to a trustworthy network of confidentiality, integrity, and availability. For machine-machine processing without user-created malware, network engineering analysis is mostly one of confidence. Machine trust is also important to enterprise systems [40].

Since we seek to understand self-confidence as a URREF criterion, explorations included human processing and the human as a data source as shown in Figure 3.

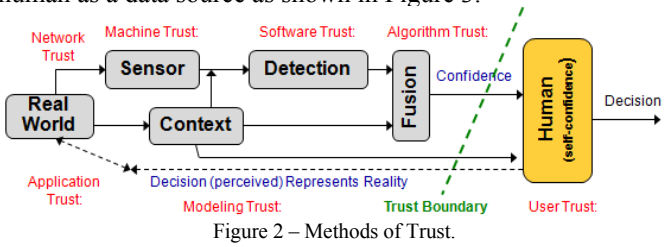


Figure 2 – Methods of Trust.

In the figure, multiple forms of trust are shown and related to the processing steps. Starting from the real world, data is placed on the network from which a machine (or sensor) processes the data. With context, a detection assessment is made for such domains as image, text, or cyber processing. That detection is the fused with the context information. For example, a detection of an object in an image is fused with contextual road information. The detection confidence is assessed and made available to the user with the context. The green line is the human-machine trust boundary as the human can look at the machine results or process the data themselves for given a level of confidence and render a decision. The dotted line then is a human assessment of whether or not the information presented represents reality and could be trusted.

Note, if the human is the only sensor source, then he/she is looking at data and making a decision. Their self-confidence could be based on the machine results from which they factor in many types of trust. For example, context, as related to the real world (see Figure 1), provides a validation of the machine (network to algorithm trust as a measure of confidence), while at the same time understanding the situation to determine if the information fusion analysis is providing meaningful and useful information towards the application of interest. Together a trusted decision is rendered based on the many factors.

Included in Figure 3 are many forms of trust in the analysis all of which can lead to confidence in the decision:

Trust	Processing	Example
Network	Data put on a network	Assessment of data timeliness and lost packets
Machine	Sensor transformation	Calibration of cameras for image content
Software	Information management	Getting the correct data from a data base (e.g., a priori data)
Algorithm	Fusion method	Target tracking and classification results
Modeling	State models	Kinematic and target recognition models (e.g., training data)
Application	Situation of interest	Analysis over the correct area (e.g. target moving on a road)
User	Situation awareness	Use of cultural and behavior (e.g. assume big cars move on roads).

The self-confidence of the user analysis includes working with data, networks, and machines. The URREF ontology must account for trust over human-machine decisions for confidence analysis. To further explore how URREF is aligned with trust, we must look at self-confidence.

B. Self-Confidence

Statistically speaking, the machine decision-making accuracy is based on the data available, the model chosen, and the estimation uncertainty associated with the measured data. Given the above analysis, we could start to derive self-confidence for machine fusion operations based on the literature in human self-confidence.

Self-confidence is the socio-psychological concept related to self-assuredness in one's personal judgment and ability. As an example, researchers are often called to review papers and after their review asked to give a quality rating of their own review based on their understanding of the subject, expertise, and experience. In another example, a person might be asked to identify an object in an image with a certain rating {unlikely, possible, probable, confirm} from which then they could determine self-confidence based on their answer. Thus, there is a need to assess "self-confidence" in relation to "confidence", which is linked to uncertainty measures of *trust*.

C. Accuracy and Precision

Self confidence is strongly related to both *precision* and *accuracy*. A source can be self confident in both the precision of its generated data (consistency or variability in its reports - such as reported variance) as well as the accuracy of its reports (the reported bias or the reported distance of the mean value of the generated data form true value). In other words, to make sense of the term the self-confidence of a source, the data encapsulate a combination of precision and accuracy. A distinction is made between precision and accuracy reported by the machine (such as the estimated mean and variance at the output of the Kalman filter) and the actual precision and accuracy of the data emanating from the source. The URREF ontology categorizes accuracy, precision, and self-confidence as types of criteria to evaluate data [1].

Statistical methods of uncertainty analysis from measurement systems include accuracy and precision, shown in Figure 4. The use of distance metrics (accuracy) and precision metrics (standard deviations) help to analyze whether the measurement is calibrated and repeatable. We would desire the same analysis for human semantic analysis with precise meanings, consistent understanding, and accurate terminology.

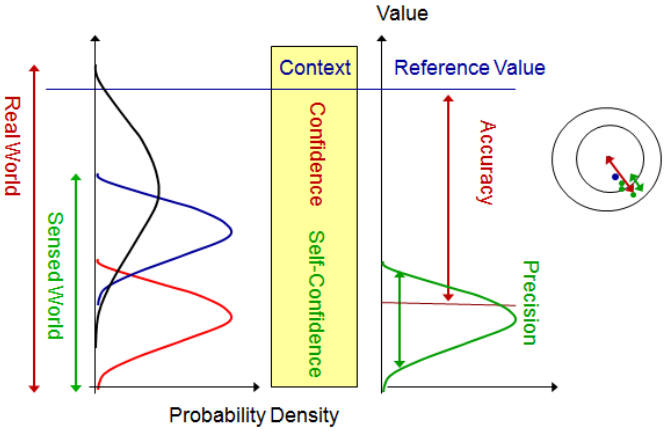


Figure 3 – Uncertainty as a function of accuracy and precision.

Human Confidence-Accuracy: Traditionally known as the confidence-accuracy (CA) relationship, the assumption is that as one's confidence increases so does their level of accuracy which is affected by memory, consistency, and ability [41]. Issues include absolute versus relative assessment, feedback, and performance.

The *confidence-accuracy* relationship was shown to be a by-product of the consistency-correctness relationship: It is positive because the answers that are consistently chosen are generally correct, but negative when the wrong answers tend to be favored. The overconfidence bias stems from the reliability-validity discrepancy: Confidence monitors reliability (or self-consistency), but its accuracy is evaluated in calibration studies against correctness. Also, the response

speed is a frugal cue for self-consistency and depends on the validity of self-consistency in predicting performance [42].

Koriat [42] explains that Sensing tasks are dominated by Thurstonian uncertainty (local rank ordering with stochastic noise) within an individual and exhibit an under-confidence bias. However, general knowledge tasks are dominated by Brunswikian Uncertainty (global probabilistic model from limited sample sets to infer general knowledge [43]) that supports inter-person ecological relations.

Consistency is then the repeatability of the information, which should imply no conflicts in decision-making. We can use the proportional conflict redistribution (PCR6) to get a measure of repeated consistency such that favored wrong answers are corrected in confidence analysis [44]. PCR6 is more general and efficient than PCR5 when combining more than two sources altogether. Moreover, PCR6 has been proved compatible with frequency probabilities when working with binary BBA's, whereas PCR5 and DS are not compatible with frequency probabilities [44].

Self-confidence could be measured with a Receiver Operating Characteristic (ROC) curve as once a decision can be made, we can then assess its impact on confidence. A low self-confidence would lead to chance, and a high self-confidence would remain to the left on the ROC.

III. Self-Confidence

Signal detection theory provides a measure of confidence in decision making that by assuming a limited hypothesis set is actually a measure of self-confidence. One classic example is Wald's Sequential Probability Ratio Test (SPRT) [45]. Assuming evidence is sampled at discrete time intervals, then the human or cognitive agent compares the conditional probabilities $x(t + \Delta t)$ for two hypothesis H_j ($j = 1, 2$). Using then SPRT, then

$$y(t) = h[x(t)] = \mathrm{LN}\left[\frac{f_1[x(t)]}{f_2[x(t)]}\right] \quad (1)$$

If $y(t) > 0$, then evidence supports H_1, and if $y(t) < 0$, then H_2 is more likely. As time accumulates for decision making, there is an aggregation of the log likelihood ratios:

$$L(t + \Delta t) = L(t) + \mathrm{LN}\left[\frac{f_1[x(t + \Delta t)]}{f_2[x(t + \Delta t)]}\right] \quad (2)$$

where, for a stochastic system $\mathrm{L(t)} \sim \mathcal{N}(\mu(t), \sigma^2(t))$. Eq (2) can be written in Bayesian log odds:

$$\mathrm{LN}\left[\frac{p(H_1 \mid D)}{p(H_2 \mid D)}\right] = \sum_t \mathrm{LN}\left[\frac{f_1[x(t)]}{f_2[x(t)]}\right] + \mathrm{LN}\left[\frac{p(H_1)}{p(H_2)}\right] \quad (3)$$

One then collects information to make a decision such that $-\theta_2 < L(t) < \theta_1$. The chosen threshold is then a measure of a decision, which can be conservative or aggressive for the case of a human agent [46]. Figure 5 shows the case in which evidence is accumulated and a decision is made with associated standard boundaries for semantic decision making. Also in Figure 5 we related decision boundaries for semantic confidence classification [47].

It is noted that a choice in time is not just the product of the current analysis, but the accumulated evidence. For example, in Figure 5, we see that the signal is moving between semantic boundaries from doubtful to probable, with an associated measure label of possible. Given the history, then the decision maker could be self-confident in the current measurement given their perception of the entire processing of machine decision making measures for each time.

A Piercian hypothesis [48] implies confidence is a multiplicative function of the quantity of the information needed to make a decision (θ, or the distance traveled by the diffusion process) and the quality of the information (δ, or the rate of evidence accumulation in the diffusion process) accumulated in Dynamic Signal Detection [48]. Without bias, the authors of [48] show that:

$$\overline{conf\,(self)} = \beta \cdot \left(\frac{1}{2}\right) \mathrm{LN}\left[\frac{P(R_A \mid S_A)}{P(R_B \mid S_A)}\right] = \frac{\delta\theta}{\sigma^2} \quad (4)$$

where β is a scaling parameter. A decision, θ, is related to a response (R) of detection to a stimuli (S). Given the ability to model self-confidence as a measure of precision, we extend the methodology using subjective-logic and DSmT [44] for robust decision making.

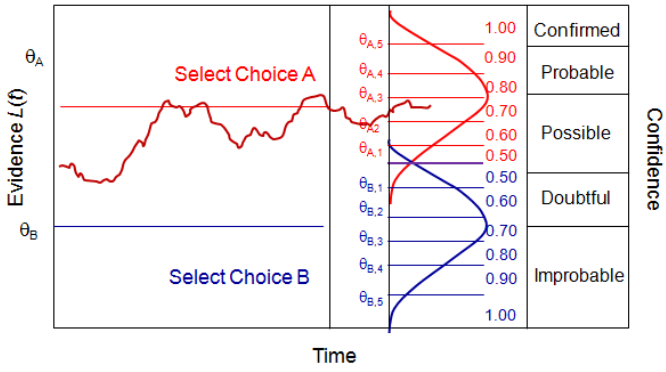


Figure 4 – Evidence Accumulation for Decision Confidence.

IV. Subjective Opinons

Subjective opinions [49] are special cases of belief functions as they correspond to bba defined on 2D frames of type $\theta = \{A, \neg A\}$ assuming Shafer's model or DSmT. Subject opinions lend themselves to simple mathematical expressions of fusion models. We therefore use the opinion representation for describing the various fusion models, but the expressions can easily be mapped to traditional belief functions.

A subjective opinion expresses belief about statements in a frame. Let X be a frame of cardinality κ. An opinion distributes belief mass over the reduced powerset $R(X)$ of cardinality κ. The reduced powerset $R(X)$ is defined as:

$$R(X) = P(X) \setminus \{X, \varnothing\}, \quad (5)$$

where $P(X) = 2^X$ denotes the powerset of X. All proper subsets of X are elements of $R(X)$, but the frame $\{X\}$ and empty set $\{\varnothing\}$ are not elements of $R(X)$.

Let $\vec{b}_X$ be a belief vector over the elements of $R(X)$, u_X be the complementary uncertainty mass, and $\vec{a}$ be a base rate vector over X. Whenever relevant, a superscript such as A denotes the opinion owner. Then a subjective opinion ω_X^A is the composite function expressed as:

$$\omega_X^A = (\vec{b}_X, u_X, \vec{a}_X). \quad (6)$$

The attribute A is thus the belief source, and X is the target frame. The belief, uncertainty and base rate parameters satisfy the following additivity constraints.

- Belief additivity:

$$u_X + \sum_{x_i \in R(X)} \vec{b}_X(x_i) = 1, \qquad \text{where } x \in R(X) \quad (7)$$

- Base rate additivity:

$$\sum_{i=1}^{k} \vec{a}_X(x_i) = 1, \qquad \text{where } x \in X \quad (8)$$

The belief vector $\vec{b}_X$ has $\kappa = (2^k - 2)$ parameters, whereas the base rate vector $\vec{a}_X$ only has k parameters. The uncertainty parameter u_X is a simple scalar. A general opinion thus contains $(2^k + k - 1)$ parameters. However, given that Eq.(7) and Eq.(8) remove one degree of freedom each, opinions over a frame of cardinality k only have $(2^k + k - 3)$ degrees of freedom. The probability projection of hyper opinions is the vector denoted as $\vec{E}_X$:

$$\vec{E}_X = \sum_{x_j \in R(X)} \vec{a}_X(x_i \mid x_j)\ \vec{b}_X(x_j) + \vec{a}_X(x_i)\, u_X, \ \forall x_i \in R(X) \quad (9)$$

$$\text{where} \quad \vec{a}_X(x_i \mid x_j) = \frac{\vec{a}_X(x_i \cap x_j)}{\vec{a}_X(x_j)}, \qquad \forall x_i, x_j \subset X. \quad (10)$$

denotes relative base rate, i.e. the base rate of subset x_i relative to the base rate of (partially) overlapping subset x_j.

General opinions are also called *hyper opinions*. A *multinomial opinion* is when belief mass only applies to singleton statements in the frame. A *binomial opinion* is when the frame is binary. A *dogmatic opinion* is an opinion without uncertainty, i.e. where $u = 0$. A *vacuous opinion* is an opinion that only contains uncertainty, i.e. where $u = 1$. Likewise, we can make the case that confidence in the opinion is biased by a subjective opinion of the *source self-confidence*. Thus, self-confidence is $SC_U = 1$ owing to rank-order decision-making on a subset of the world, and the lack of self-confidence is $SC_U = 0$; where:

$$SC_U(\omega_X^A) = \vec{a}_X[1 - u_X] \quad (11)$$

$$\text{and} \qquad \omega_X^A \leftarrow SC_U(\omega_X^A) \bullet \vec{b}_X \quad (12)$$

Equivalent probabilistic representations of opinions, e.g. as a Beta pdf (probability density function) in case of binomial opinions, as a Dirichlet pdf in case of multinomial opinions, or as a hyper Dirichlet pdf in case of hyper opinions offer an alternative interpretation of subjective opinions in terms of traditional statistics [6].

Cumulative Fusion:
The cumulative fusion rule is equivalent to *a posteriori* updating of Dirichlet distributions. Its derivation is based on the bijective mapping between the belief and evidence notations described in [6].

The symbol "◊" denotes the cumulative fusion of two observers A and B into a single imaginary observer $A \diamond B$.

Let ω^A and ω^B be opinions respectively held by agents A and B over the same frame X of cardinality k with reduced powerset $R(X)$ of cardinality κ. Let $\omega^{A \diamond B}$ be the opinion where:

CASE I: For $u^A \neq 0 \ \vee \ u^B \neq 0$ (with Confidence)

$$\begin{cases} b^{A \diamond B}(x_i) = \dfrac{b^A(x_i)\, u^B + b^B(x_i)\, u^A}{u^A + u^B - u^A u^B} \\ u^{A \diamond B} = \dfrac{u^A u^B}{u^A + u^B - u^A u^B} \end{cases} \quad (13)$$

CASE II: For $u^A = 0 \ \vee \ u^B \neq 0$ (without Confidence)

$$\begin{cases} b^{A \diamond B}(x_i) = \gamma^A b^A(x_i) + \gamma^B b^B(x_i) \\ u^{A \diamond B} = 0 \end{cases} \quad (14)$$

$$\text{where:} \quad \begin{cases} \gamma^A = \lim_{u^A \to 0;\, u^B \to 0} \dfrac{u^B}{u^A + u^B} \\ \gamma^B = \lim_{u^A \to 0;\, u^B \to 0} \dfrac{u^A}{u^A + u^B} \end{cases}$$

Note: the case without confidence averages the results from self-confidence reports which weights effectively both the same. Confidence allows the user to weight the self-confidence of the reports based on the Brunswikian uncertainty about the world knowledge.

Then $\omega^{A \diamond B}$ is the cumulatively fused opinion of ω^A and ω^B, representing the combination of independent opinions of A and B. By using the symbol '$\oplus$' to designate this belief operator, cumulative fusion is expressed as:

$$\text{Cumulative Belief Fusion:} \quad \omega_X^{A \diamond B} = \omega_X^A \oplus \omega_X^B \quad (15)$$

The cumulative fusion operator is commutative, associative and non-idempotent. In Eq.(15), the associativity depends on the preservation of relative weights of intermediate results through the weight variable γ, in which case the cumulative rule is equivalent to the weighted average of probabilities.

V. Results

Assume we have two agent opinion makers ω^A and ω^B, who each make a decision for network security [50]. Let ω^A be a machine *A*lgorithm and let ω^B come from a human *B*eing. After reporting their opinion, ω^B is asked for their self-confidence. The result modifies their belief $\vec{b}_X$, such that the cumulative belief fusion product is a weighted function of

their self-confidence (source) over their confidence (data). Figure 6 provides a perspective of the analysis.

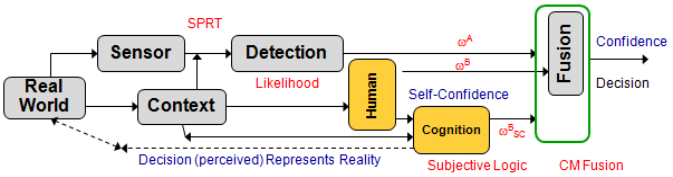


Figure 5 – Analysis With and Without Self Confidence.

For many situations, a machine can process large amounts of data, while a human agent can only comprehend a subset of the data. Thus, a machine processes the data as outputs to a user. The interaction with the user is continually updated and a decision from the user is required. For situations in which the user has more time (forensics), then his/her self-confidence in the data would be high. For quick decisions, an observe-orient-decide-act (OODA) decision might be required [51] which reduces self-confidence. We seek methods of the latter as uncertainty is higher in rapid decision making which is a subset of problems in the Dynamic Data-Driven Application Systems (DDDAS) paradigm [52, 53].

For the analysis, we have two opinion makers (machine and man). Using signal detection theory, their individual measures of analysis provide a likelihood function. We then fuse the results with confusion matrix fusion [54] as a method of combination using Bayesian, Dempster-Shafer, or DSmT results [55]. We utilize two cases in which there is a high and low-confident observer (Case 1) and then the situation in which both have comparable analysis (Case 2). With two highly self-confident observers (Case 2), the results are similar to one of the observers which could be used for opinion validity. However, the user could be looking at the results and further analyzing context to provide a more appropriate analysis of their decision (e.g., based on culture, data completeness, etc). Using subjective logic the human being could modify their opinion, ω^B_{SC}, which results in a larger value (e.g., know something) or lower value (e.g., recognize limitation of analysis).

We assume that if the user provides no assessment of self-confidence, we provide equal weight to the results (average fusion). On the other hand, if a machine provides a measure of confidence, it could be derived from the dynamic-data, which we don't simulate here.

Example (*High self-confidence with low self-confidence*)
Assume that we have a highly self-confident opinion maker, ω^A, that includes many sources and reliable analysis. On the other hand, we have a low-confident opinion maker, ω^B, who is making a decision. When making their decision, ω^B is guessing or almost chance, assuming that context provides pragmatic understanding of the world events.

In Figure 7, there are two opinion makers, the red curve of a human agent suggesting that the result is "improbable," while the more self-confident is in blue reporting "probable". The fused result, shown in green, using self-confidence better reflects the true state; versus the average fusion of the opinion makers shown in magenta. The key issue is that self-confidence can help weight evidence.

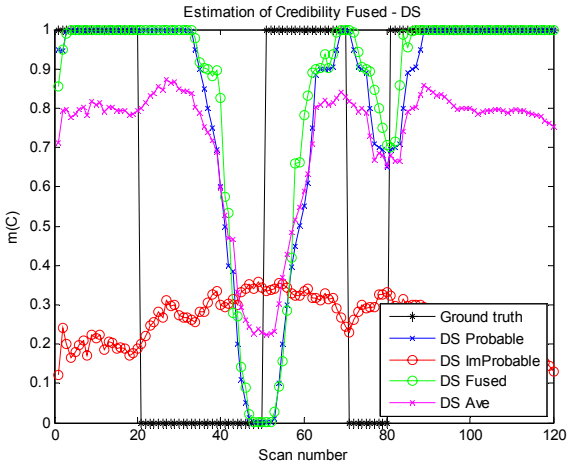


Figure 6 – DS With (Fused) and With-out (Ave) Self-Confidence.

Exploring DSmT [44], using the proportional conflict redistribution rule (PCR6[1]), we also see in Figure 8 an improvement in the belief confidence when self-confidence is accounted for.

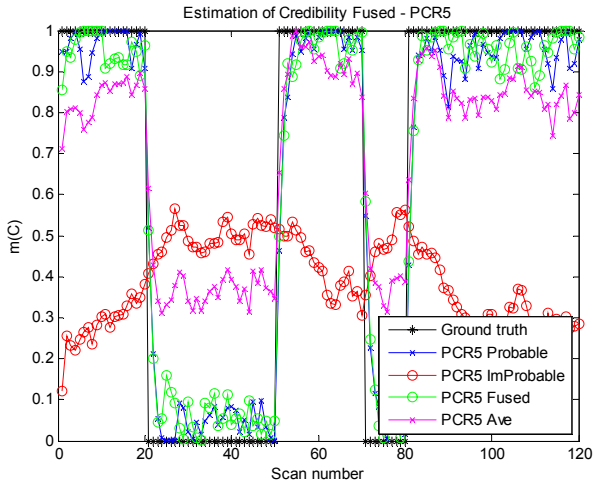


Figure 7 – PCR6 With (Fused) and With-out (Ave) Self-Confidence.

VI. CONCLUSIONS

In this paper, we assessed self-confidence as a criterion in the URREF. Self-confidence is typically associated with a source and relates a subjective quality on the rendering of their beliefs over data. For stochastic observations, we use the SPRT in a self-confidence analysis. However, to get the case of partial information, we use subjective logic for decision-makers. We demonstrated that the PCR6 is superior to DS for decision for a scenario in which a high self-confident observer opinion is fused with a low self-confident observer. Ultimately it is the user trust in the data they have available and opinions towards self-confidence; whereas a machine only reports confidence.

Further directions include using the analysis with real operators doing intelligence analysis over data and associating semantic boundaries to their subjective decision-making.

[1] In the scenario, we used sequential fusion of two sources and because of this, PCR5=PCR6, i.e. when combining 2 sources only PCR5 coincides with PCR6.

ACKNOWLEDGEMENTS

Erik Blasch was supported under an AFOSR grant in Dynamic Data-Driven Applications Systems (DDDAS) grant. Paulo Costa was partially supported by the Army Research Office, under grant W911NF-11-1-0176.

REFERENCES

[1] P. C. G. Costa, K. B. Laskey, E. Blasch, and A-L. Jousselme, "Towards Unbiased Evaluation of Uncertainty Reasoning: The URREF Ontology," *Int. Conf. on Info Fusion*, 2012.

[2] E. Blasch, K. B. Laskey, A-L. Jousselme, V. Dragos, P. C. G. Costa, and J. Dezert, "URREF Reliability versus Credibility in Information Fusion (STANAG 2511)," *Int'l Conf. on Info Fusion*, 2013.

[3] B. Kahler and E. Blasch, "Predicted Radar/Optical Feature Fusion Gains for Target Identification," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON)*, 2010.

[4] E. Blasch, I. Kadar, J. Salerno, M. M. Kokar, S. Das, *et. al.*, "Issues and Challenges in Situation Assessment (Level 2 Fusion)," *J. of Advances in Information Fusion,* Vol. 1, No. 2, pp. 122 - 139, Dec. 2006.

[5] R. Hummel and M. S. Landy, "A statistical viewpoint on the theory of evidence," *IEEE Trams. Pattern Analysis and Machine Intelligence*, Vol. 10 (2), pp. 235-247, 1988.

[6] A. JØsang, P. C. G. Costa, E. Blasch, "Determining Model Correctness for Situations of Belief Fusion," *Int'l Conf. on Info Fusion*, 2013.

[7] J. Dezert, J.-M. Tacnet, M. Batton-Hubert, F. Smarandache, "Multi-Criteria decision making based on DSmT-AHP," *Workshop on theory of Belief Functions*, 2010.

[8] P. Smets, R. Kennes, "The transferable belief model," *J. Artificial Intelligence*, Vol. 66, pp. 191--243, 1994.

[9] J. Dezert, D. Han, Z-G Liu, J-M Tacnet, "Hierarchical DSmP transformation for decision-making under uncertainty," *Int. Conf. on Information Fusion*, 2012.

[10] E. Blasch & S. Plano, "Level 5: User Refinement to aid the Fusion Process," *Proc. of SPIE,* Vol. 5099, 2003.

[11] E. Blasch and S. Plano, "DFIG Level 5 (User Refinement) issues supporting Situational Assessment Reasoning," *Int. Conf. on Information Fusion*, 2005.

[12] C. Yang and E. Blasch, "Fusion of Tracks with Road Constraints," *J. of. Advances in Information Fusion,* Vol. 3, No. 1, 14-32, June 2008.

[13] G. Chen, D. Shen, C. Kwan, J. Cruz, *et al*., "Game Theoretic Approach to Threat Prediction and Situation Awareness," *Journal of Advances in Information Fusion,* Vol. 2, No. 1, 1-14, June 2007.

[14] E. P. Blasch, E. Bosse, and D. Lambert, *High-Level Information Fusion Management and Systems Design*, Artech House, Norwood, MA, 2012.

[15] P. Foo, G. Ng, "High-Level Information Fusion: An Overview," *J. Adv. Information Fusion*, Vol. 8 (1), June 2013.

[16] S. K. Rogers, M. Kabrisky, K. Bauer, and M. Oxley., "Computing Machinery and Intelligence Amplification," in *Computational Intelligence: The Experts Speak,* ed. David B. Fogel and Charles J. Robinson, Piscataway, NJ: IEEE Press , 25-44, 2003.

[17] S. K. Rogers *et al*., "The life and death of ATR/sensor fusion and the hope for resurrection," *Proc. of SPIE*, Vol. 6967, 2008.

[18] E. Blasch and P. Hanselman, "Information Fusion for Information Superiority," *IEEE Nati'l Aerospace and Electronics Conf.,,* 2000.

[19] E. P. Blasch "Assembling a distributed fused Information-based Human-Computer Cognitive Decision Making Tool," *IEEE Aerospace and Electronic Systems Magazine,* Vol. 15, No. 5, pp. 11-17, May 2000.

[20] E. Blasch and S. Plano, "JDL Level 5 Fusion model 'user refinement' issues and applications in group Tracking," *Proc. SPIE,* 4729, 2002.

[21] A. Panasyuk, *et al*., "Extraction of Semantic Activities from Twitter Data," *Semantic Tech. for Intelligence, Defense, and Security*, 2013.

[22] C. Yang and E. Blasch, "Kalman Filtering with Nonlinear State Constraints," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 45, No. 1, 70-84, Jan. 2009.

[23] A. JØsang, "The Consensus operator for combining beliefs," *Artificial Intelligence*, 2002.

[24] P. C. G. Costa, K . B. Laskey, "PR-Owl: A framework for probabilistic ontologies," *Frontiers in Artificial Intelligence*, 2006.

[25] E. Blasch, P. Valin, and E. Bossé, "Measures of Effectiveness for High-Level Fusion," *Int'l Conf. on Info Fusion*, 2010.

[26] E. Blasch, P. C. G. Costa, K. B. Laskey, D. Stampouli, G. W. Ng, J. Schubert, R. Nagi, and P Valin, "Issues of Uncertainty Analysis in High-Level Information Fusion – Fusion2012 Panel Discussion," *Int. Conf. on Info Fusion*, 2012.

[27] P. C. G. Costa, E. P. Blasch, K. B. Laskey, S. Andler, J. Dezert, A-L. Jousselme, and G. Powell, "Uncertainty Evaluation: Current Status and Major Challenges – Fusion2012 Panel Discussion," *Int. Conf. on Info Fusion*, 2012.

[28] E. Blasch, M. Pribilski, *et al*., "Fusion Metrics for Dynamic Situation Analysis," *Proc. of SPIE,* Vol. 5429, 2004.

[29] E. Blasch, I. Kadar, J. Salerno, M. M. Kokar, *et al*., "Issues and challenges of knowledge representation and reasoning methods in situation assessment (Level 2 Fusion)," *Proc. of SPIE,* Vol. 6235, 2006.

[30] T. Schuck, and E. P. Blasch, "Description of the Choquet Integral for Tactical Knowledge Representation," *Int'l. Conf. on Info. Fusion*, 2010.

[31] J. Debenham, "A Logic of Knowledge Integrity," University of Technology, Sydney, NSW 2007.

[32] J. D. Lee, N. Moray, "Trust, self-confidence, and operators' adaptation to automation," *International Journal of Human-Computer Studies*, 1994.

[33] E. Blasch, Y Al-Nashif, and S. Hariri, "Static versus Dynamic Data Information Fusion analysis using DDDAS for Cyber Trust," *International Conference on Computational Science*, 2014.

[34] H. Aras, C. Beckstein, *et al*., "Uncertainty and Trust," Univ. of Illinois, Dagstuhl Seminar Proceedings 08421, *Uncertainty Management in Information Systems*, 2009.

[35] A, Jøsang, S. L. Presti, "Analyzing the relationship between risk and trust," International *Conference on Trust Management*, 2004.

[36] D. Wang, T. Abdelzaher, and L. Kaplan, "On truth discovery in social sensing: A maximum likelihood estimation approach," UIUC Technical Report, 2011.

[37] D. Wang, T. Abdelzaher, L. Kaplan, and C. Aggarwal, "On quantifying the accuracy of maximum likelihood estimation of participant reliability in social sensing," in *DMSN*, pages 7–12, 2011.

[38] E. Onal, J. Schaffer, J. O'Donovan, et al., "Decision-Making in Abstract Trust Games: A User Interface Perspective," *IEEE CogSIMA*, 2014.

[39] D. Shen, G. Chen, *et al*., "A Trust-based Sensor Allocation Algorithm in Cooperative Space Search Problems," *Proc. SPIE* , Vol. 8044, 2011.

[40] E. Blasch, O. Kessler, *et al*., "Information Fusion Management and Enterprise Processing." *IEEE Nat. Aerospace and Elect. Conf.*, 2012.

[41] K Krug, "The relationship between confidence and accuracy: current thoughts of the literature and a new area of research," *Applied Psychology in Criminal Justice*, 3(1), 2007.

[42] A. Koriat, "The Self-Consistency Model of Subjective Confidence," Psychological Review, 119(1)' 80-113, 2012.

[43] G. Gigerenzer, U. Hoffrage, and H. Kleinbolting, "Probabilistic Mental Models: A Brunswikian Theory of confidence," *Psychological Review*, 98(4), 506-538, 1991.

[44] F. Smarandache, J. Dezert, "On the consistency of PCR6 with the averaging rule and its application to probability estimation," *Int'l Conf. on Information Fusion*, 2013.

[45] A. Wald, "Sequential Tests of Statistical Hypotheses," *Annals of Mathematical Statistics* 16 (2): 117–186, June 1945.

[46] E. Blasch and R. Broussard, "Physiologically Motivated Computational Visual Target Recognition Beta Selection," *Proc. SPIE,* Vol. 4055, Orlando, FL, April 2000.

[47] E. Blasch, R. Breton, and P. Valin, "Information Fusion Measures of Effectiveness (MOE) for Decision Support," *Proc. SPIE*, 8050, 2011.

[48] T. J. Pleskac and J. B. Busemeyer, "Two-stage Dynamic Signal Detection: A Theory of Choice, Decision Time, and Confidence," Psychological Review, Vol. 117, No. 3, 864-901, 2010.

[49] A. Jøsang and R. Hankin, "Interpretation and Fusion of Hyper Opinions in Subjective Logic," *Int'l Conf. on Information Fusion*, 2012.

[50] A. Jøsang, "Prospectives for Modeling Trust in Information Security," *Information Security and Privacy*, 1997.

[51] E. Blasch, R. Breton, *et al*., "User Information Fusion Decision Making Analysis with the C-OODA Model," *Int. Conf. on Info Fusion*, 2011.

[52] E. Blasch, G. Seetharaman, *et al*., "Dynamic Data Driven Applications Systems (DDDAS) modeling for Automatic Target Recognition," *Proc. SPIE*, Vol. 8744, 2013.

[53] E. Blasch, G. Seetharaman, *et al*., "Dynamic Data Driven Applications System concept for Information Fusion," *International Conference on Computational Science*, 2013. (ICCS13), *Procedia Computer Science*, Vol. 18, Pages 1999-2007, 2013.

[54] B. Kahler and E. Blasch, "Decision-Level Fusion Performance Improvement from Enhanced HRR Radar Clutter Suppression," *J. of. Advances in Information Fusion,* Vol. 6, No. 2, Dec. 2011.

[55] E. Blasch, J. Dezert, B Pannetier, "Overview of Dempster-Shafer and Belief Function Tracking Methods," *Proc. SPIE*, Vol. 8745, 2013.

Evaluations of Evidence Combination Rules in Terms of Statistical Sensitivity and Divergence

Deqiang Han
Jean Dezert
Yi Yang

Originally published as Han D., Dezert J., Yang Y., *Evaluations on Evidence Combination Rules in Terms of Statistical Sensitivity and Divergence,* in Proc. of Fusion 2014 Int Conf onInformation Fusion, Salamanca, Spain, July 7-10, 2014, and reprinted with permission.

***Abstract*—The theory of belief functions is one of the most important tools in information fusion and uncertainty reasoning. Dempster's rule of combination and its related modified versions are used to combine independent pieces of evidence. However, until now there is still no solid evaluation criteria and methods for these combination rules. In this paper, we look on the evidence combination as a procedure of estimation and then we propose a set of criteria to evaluate the sensitivity and divergence of different combination rules by using for reference the mean square error (MSE), the bias and the variance. Numerical examples and simulations are used to illustrate our proposed evaluation criteria. Related analyses are also provided.**

***Keywords*—belief functions; evidence combination; evaluation; sensitivity; divergence.**

I. Introduction

The theory of belief functions, also called Dempster-Shafer evidence Theory(DST) [1], is one of the most important theories and methods in information fusion and uncertainty reasoning. It can distinguish 'unknown' and 'imprecision' and propose a way to fuse or combine different pieces of evidence by using the commutative and associative Dempster's rule of combination.

Dempster's rule of combination can bring counter-intuitive combination results in some cases [2], [3], so there have emerged several improved and modified alternative evidence combination rules, where counter-intuitive behaviors are imputed to the combination rule itself, especially the way to deal with the conflicting mass assignments. The representative works include Yager's rule [4], Florea's robust combination rule (RCR) [5], disjunctive rule [6], Dubois and Prade's rule [7], proportional conflict redistribution rule (PCR) [8], and mean rule [9], etc.

As aforementioned, several combination rules are available including Dempster's rule and its alternatives. Then, how to evaluate them? This is crucial for the practical use of the combination rules. The qualitative criterion is that the combination results should be intuitive and rational [10]. Up to now, there is still no solid performance evaluation approaches for combination rules, especially for establishing quantitative criteria. In this paper, we propose to interprept the evidence combination as a procedure of estimation [11]; therefore, a combination rule is regarded as an estimator. So, we define some statistical criteria on sensitivity and divergence for the different combination rules by using for reference the idea of Mean Square Error (MSE) and its decomposition in estimation. By adding small errors to the original pieces of evidence (i.e., the "input" of the "estimator"), we check the mean square error, the variance, and the bias of the combination result ("output" of the estimator) caused by adding some noise to describe the sensitivity and divergence of the given combination rule. Distance of evidence [12] is used in our work to define the variance, the bias and other related criteria. Simulation results are provided to illustrate our proposed evaluation approaches. Dempster's rule and major available alternative rules are evaluated and analyzed using the new evaluation approaches.

II. Basics of DST

Dempster-Shafer evidence theory (DST) [1] has been developed by Shafer in 1976 based on previous works of Dempster. In evidence theory, the elements in frame of discernment (FOD) Θ are mutually exclusive and exhaustive. Define $m : 2^\Theta \to [0, 1]$ as a basic belief assignment (BBA, also called mass function) satisfying:

$$\sum_{A\in 2^\Theta} m(A) = 1, \quad m(\emptyset) = 0 \tag{1}$$

If $m(A) > 0$, A is called a focal element. In DST, two reliable independent bodies of evidence (BOEs) $m_1(\cdot)$ and $m_2(\cdot)$ are combined using Dempster's rule of combination as follows. $\forall A \in 2^\Theta$:

$$m(A) = \begin{cases} 0, A = \emptyset \\ \frac{\sum_{A_i\cap B_j=A} m_1(A_i)m_2(B_j)}{1-K}, A \neq \emptyset \end{cases} \tag{2}$$

where

$$K = \sum\nolimits_{A_i\cap B_j=\emptyset} m_1(A_i)m_2(B_j) \tag{3}$$

represents the total conflicting or contradictory mass assignments. Obviously, from Eq. (2), it can be verified that Dempster's rule is both commutative and associative. For Dempster's rule of combination, the conflicting mass assignments are discarded through a classical normalization step.

As firstly pointed out by Zadeh [2], Dempster's rule has been criticized for its counter-intuitive behaviors[1]. DST's validity has also been argued [3]. There have emerged several alternatives of evidence combination rules aiming to suppress the counter-intuitive behaviors of classical Dempster's rule. See [8] for details.

To measure the dissimilarity between different BBAs, the distance of evidence can be used. Jousselme's distance [13] is one of the most commonly used distance of evidence, which is defined as

$$d_J(m_1, m_2) = \sqrt{\frac{1}{2} \cdot (m_1 - m_2)^T \mathbf{Jac} \ (m_1 - m_2)} \tag{4}$$

where the element $J_{ij} \triangleq \mathbf{Jac}(A_i, B_j)$ of Jaccard's weighting matrix $\mathbf{Jac}$ is defined as

$$\mathbf{Jac}(A_i, B_j) = \frac{|A_i \cap B_j|}{|A_i \cup B_j|} \tag{5}$$

There are also other types of distance of evidence [12], [14]. We choose to use Jousselme's distance of evidence in this paper, because it has been proved to be a strict distance metric [15].

III. Some major alternative combination rules

In this section, some major combination rules in evidence theory other than Dempster's rule are briefly introduced. For all $A \in 2^\Theta$

1) Yager's rule [4]:

$$\begin{cases} m(\emptyset) = 0 \\ m_{Yager}(A) = \sum\limits_{B_i \cap C_j = A \neq \emptyset} m_1(B_i) m_2(C_j) \\ m(\Theta) = m_1(\Theta) m_2(\Theta) + \sum\limits_{B \cap C = \emptyset} m_1(B_i) m_2(C_j) \end{cases} \tag{6}$$

In Yager's rule, the conflict mass assignments are assigned to the total set of the FOD Θ.

2) Disjunctive rule [6]:

$$\begin{cases} m(\emptyset) = 0 \\ m_{Dis}(A) = \sum\limits_{B_i \cup C_j = A} m_1(B_i) m_2(C_j) \end{cases} \tag{7}$$

This rule reflects the disjunctive consensus.

3) Dubois & Prade's rule (D&P rule) [7]:

$$\begin{cases} m(\emptyset) = 0 \\ m_{DP}(A) = \sum\limits_{B_i \cap C_j = A \neq \emptyset} m_1(B_i) m_2(C_j) \\ \quad + \sum\limits_{B_i \cap C_j = \emptyset, B_i \cup C_j = A} m_1(B_i) m_2(C_j) \end{cases} \tag{8}$$

This rule admits that the two sources are reliable when they are not in conflict, but only one of them is right when a conflict occurs.

[1]According to the viewpoint of proponents of Dempster's rule, the counter-intuitive behavior is imputed to the sensors, the data or the BOEs obtained from different sources, but not to Dempster's rule itself.

4) Robust Combination Rule (RCR, or Florea's rule) [5]:

$$m_{RCR}(A) = \alpha(K) m_{Dis}(A) + \beta(K) m_{Conj}(A) \tag{9}$$

where m_{Dis} is the BBA obtained using the disjunctive rule, m_{Conj} is the BBA obtained using the conjunctive rule, and $\alpha(K), \beta(K)$ are the weights, which should satisfy

$$\alpha(K) + (1 - K)\beta(K) = 1 \tag{10}$$

where K is the conflict coefficient defined in Eq. (3). Robust combination rule can be considered as a weighted summation of the BBAs obtained using the disjunctive rule and the conjunctive rule, respectively.

5) PCR5 [8]: Proportional Conflict Redistribution rule 5 (PCR5) redistributes the partial conflicting mass to the elements involved. in the partial conflict, considering the canonical form of the partial conflict. PCR5 is the most mathematically exact redistribution of conflicting mass to non-empty sets following the logic of the conjunctive rule.

$$\begin{aligned} & m_{PCR5}(\emptyset) = 0 \\ & \text{and } \forall X \in 2^\Theta \setminus \{\emptyset\} \\ & m_{PCR5}(A) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \\ & \quad \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)} \right] \end{aligned} \tag{11}$$

In fact there exists another rule PCR6 that coincides with PCR5 when combining two sources, but differs from PCR5 when combining more than two sources altogether and PCR6 is considered more efficient than PCR5 because it is compatible with classical frequentist probability estimate [16].

6) Mean rule [9]:

$$m_{mean}(A) = \frac{1}{n} \sum\nolimits_{i=1}^{n} m_i(A) \tag{12}$$

By using this rule, we can find the average of the BBAs to be combined.

For the purpose of the practical use of different combination rules, the evaluation criteria are required. In the next section, the available evaluation criteria or properties of evidence combination rules are beriefly introduced.

IV. Properties of combination rules as qualitative criteria

1) Commutativity [17]: The combination of two BBAs m_1 and m_2 using some rule R does not depend on the order of the two BBA, i.e.,

$$R(m_1, m_2) = R(m_2, m_1) \tag{13}$$

All the combination rules aforementioned in Section III are commutative.

2) Associativity [17]: The combination result of multiple BBAs does not depend on the order of the BBAs to be combined. For example, when there are 3 BBAs,

$$R(R(m_1, m_2), m_3) = R(m_1, R(m_2, m_3)) \tag{14}$$

Dempster's rule and disjunctive rule are associative. The other rules introduced in Section III are not associative. The property of associativity is important to facilitate the implementation of the distributed information fusion system. But it should be noted that it is not necessarily efficient in term of quality of fusion result. Non-associative rules are able to provide better performances in general than associative rules [16].

3) Neutral impact of the vacuous belief [17]: The combination rule preserves the neutral impact of the vacuous BBA, i.e., when m_2 is $m(\Theta) = 1$,

$$R(m_1, m_2) = m_1 \tag{15}$$

All the rules aforementioned in Section III but the mean rule, satisfy this property.

These criteria are qualitative and they correspond to good (interesting) properties that a rule could satisfy. It should be noted that these "expected good" properties do not warrant that a real efficient fusion rule must absolutely satisfy them. Therefore, these properties are not enough to the evaluations of combination rules. In this paper, we propose some quantitative evaluation criteria for combination rules.

V. Statistical sensitivity and divergence of combination rules

Here, we develop a group of criteria for combination rules in terms of sensitivity and divergence. The idea of Mean Square Error (MSE) and its decomposition are used as a basic framework for such a development.

A. Mean Square Error and its decomposition

For an estimate $\hat{x}$ of the scalar estimand x, the MSE is defined as

$$\mathrm{MSE}(\hat{x}) = E[(\hat{x} - x)^2] \tag{16}$$

MSE can be decomposed as

$$\begin{aligned} \mathrm{MSE}(\hat{x}) &= E[(\hat{x} - E(\hat{x}))^2] + E[(E(\hat{x}) - x)^2] \\ &= \mathrm{Var}(\hat{x}) + (\mathrm{Bias}(\hat{x}, x))^2 \end{aligned} \tag{17}$$

The MSE is equal to the sum of the variance and the squared bias of the estimator or of the estimations. The variance can represent the divergence of the estimation results. The bias can represent the sensitivity of the estimator.

B. Criteria for statistical sensitivity and divergence

If we consider the procedure of evidence combination with a given rule as an estimator (as illustrated in Fig. 1), then we can consider the combination results as the estimations.

So, we can use for reference the MSE and its decompositions to measure the error, the variance, and the bias of the combination results based on the given combination rule. Here we attempt to design some criteria related to the sensitivity and divergence of combination rules. We use the change of the combination results after adding small noise to the original BBA to reflect the sensitivity and divergence of a combination rule. If under a given small noise, a combination rule bring out smaller variance and smaller bias, then such a rule is less divergent and less sensitive, based on which, the sensitivity and divergence of combination rules can be evaluated. The definitions of MSE, variance and bias for combination rules, and the evaluation procedure are as follows.

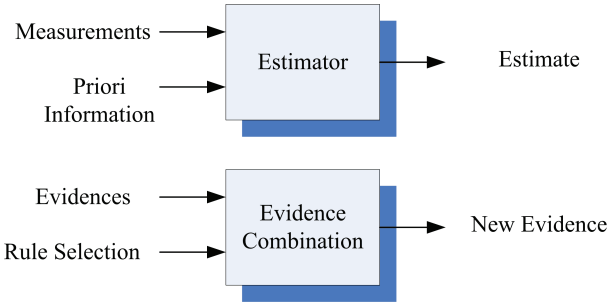


Fig. 1. Evidence combination and Estimation.

Step 1: Randomly generate a BBA m. Add random noise to m for N times, respectively. In each time, the noise is ϵ_i (small values), where $i = 1, ..., N$. The noise sequence is denoted by $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, ..., \epsilon_N]$. Here each ϵ_i is a small real number (negative or positive) close to zero. Then, we can obtain a sequence of noised BBAs as

$$\mathbf{m}' = [m_1, m_2, ..., m_N] \tag{18}$$

It should be noted that all the noised BBAs are normalized.

Step 2: Generate original combination results sequence with a combination rule R

$$\begin{aligned} \mathbf{m}_c &= [m_c^1, m_c^2, ..., m_c^N] \\ &= [R(m, m), R(m, m), ..., R(m, m)] \end{aligned} \tag{19}$$

The length of $\mathbf{m}_c$ is N.

Step 3: Generate combination results sequence by combining BBAs with noise and the original BBAs using the rule of R

$$\begin{aligned} \mathbf{m}_{cn} &= [m_{cn}^1, m_{cn}^2, ..., m_{cn}^N] \\ &= [R(m_1, m), R(m_2, m), ..., R(m_N, m)] \end{aligned} \tag{20}$$

Step 4: Calculate the MSE of $\mathbf{m}_{cn}$ as

$$\begin{aligned} \mathrm{MSE}_{\mathrm{BBA}}(\mathbf{m}_{cn}) &= \tfrac{1}{N} \sum_{i=1}^{N} \left[d_J(m_c^i, m_{cn}^i)\right]^2 \\ &= \tfrac{1}{N} \textstyle\sum_{i=1}^{N} \left[d_J(R(m, m), R(m, m_i))\right]^2 \end{aligned} \tag{21}$$

where d_J is Jousselme's distance defined in Eq. (3). $\mathrm{MSE}_{\mathrm{BBA}}$ represents the error between the original combination results and the results obtained using BBAs with noise.

We can also calculate the relative MSE by removing the effect of the noise amplitude as follows

$$\mathrm{MSE}'_{\mathrm{BBA}}(\mathbf{m}_{cn}) = \frac{\mathrm{MSE}_{\mathrm{BBA}}(\mathbf{m}_{cn})}{\|\boldsymbol{\epsilon}\|^2} \tag{22}$$

Step 5: Calculate the variance of $\mathbf{m}_{cn}$ as

$$\begin{aligned}\mathrm{Var}_{\mathrm{BBA}}(\mathbf{m}_{cn}) &= \tfrac{1}{N}\sum_{i=1}^{N}\left[d_J(m_{cn}^i, \bar{m}_{cn})\right]^2 \\ &= \tfrac{1}{N}\sum_{i=1}^{N}\left[d_J(R(m,m_i), \tfrac{1}{N}\sum_{j=1}^{N}R(m,m_j))\right]^2\end{aligned} \tag{23}$$

where $\bar{m}_{cn} = \frac{1}{N}\sum_{j=1}^{N} m_{cn}^j = \frac{1}{N}\sum_{j=1}^{N} R(m,m_j)$. $\mathrm{Var}_{\mathrm{BBA}}$ represents the fluctuations of the combination results obtained using BBAs with noise.

Then, calculate the relative variance by removing the effect of the noise amplitude as follows

$$\mathrm{Var}'_{\mathrm{BBA}}(\mathbf{m}_{cn}) = \frac{\mathrm{Var}_{\mathrm{BBA}}(\mathbf{m}_{cn})}{Var(\boldsymbol{\epsilon})} \tag{24}$$

Relative variance in fact represents the degree of amplification or reduction of the variances between and after combination.

Step 6: Calculate the bias of $\mathbf{m}_{cn}$ as

$$\begin{aligned}\mathrm{Bias}_{\mathrm{BBA}}(\mathbf{m}_{cn}) &= \sqrt{\tfrac{1}{N}\sum_{i=1}^{N}\left[d_J(m_c^i, \bar{m}_{cn})\right]^2} \\ &= \sqrt{\tfrac{1}{N}\sum_{i=1}^{N}\left[d_J(R(m,m), \tfrac{1}{N}\sum_{j=1}^{N}R(m,m_j))\right]^2} \\ &= d_J(R(m,m), \tfrac{1}{N}\sum_{j=1}^{N}R(m,m_j))\end{aligned} \tag{25}$$

$\mathrm{Bias}_{\mathrm{BBA}}$ represents the difference between the expectation of the combination results obtained using BBAs with noise and the original combination results.

Then, calculate the relative bias by removing the effect of the noise amplitude as follows

$$\mathrm{Bias}'_{\mathrm{BBA}}(\mathbf{m}_{cn}) = \frac{\mathrm{Bias}_{\mathrm{BBA}}(\mathbf{m}_{cn})}{\|\boldsymbol{\epsilon}\|} \tag{26}$$

Regenerate randomly a new original BBA m for M times. In each time, re-do Step 1 to Step 6. Based on the M groups of results, calculate the averaged $\mathrm{MSE}'_{\mathrm{BBA}}$, the averaged $\mathrm{Var}'_{\mathrm{BBA}}$, and the averaged $\mathrm{Bias}'_{\mathrm{BBA}}$. These three indices are called the statistical MSE, the statistical variance, and the statistical bias of the combination rule R. We jointly use these indices (quantitative criteria) to describe the statistical sensitivity and divergence of a given combination rule R.

Relative MSE is a comprehensive index. Larger relative MSE intuitively means larger sensitivity. However, relative MSE is insufficient to evaluate a combination rule. So we should further use its decomposition (including the relative variance and the relative bias) for a deeper analysis.

High relative bias values represent high sensitivity. it represents high degree of departure from the origin. It can reflect a given combination rule's capability of sensitive response to the changes in input evidences. It represents the "agility" of a combination rule. Moderate relative bias values are preferred, which means the balance or trade-off between the robustness and the sensitivity.

Relative variance in fact represents the degree of amplification or reduction of the variances between and after combination. In the evaluation procedure, for all the combination rule, the variance of the noise are the same (using the same noise sequence for different rules). So, high relative variance values also represent high divergence among all the combination results using a given combination rule when adding noise. Small relative variance values are preferred, which represent the high cohesion of a given combination rule.

In this work, we propose a statistical evaluation approach for evidence combination rules based on Monte-Carlo simulation. To implement the statistical evaluation of a combination rule according to the method introduced here, two problems should be resolved at first. One is the way of adding noise and the other is the way of random generation of BBA.

C. Method I for adding noise

Method I for adding noise is designed to evaluate the effect of the slight value change of the mass of the existing focal element. Suppose that m is a BBA defined on FOD Θ. First, we find the primary focal element (the focal element having the highest mass assignment) [2], i.e., the focal element A_i satisfying

$$i = \arg\max_{j, A_j \subseteq \Theta} m(A_j) \tag{27}$$

Second, add the noise ϵ to the mass assignment of the primary focal element.

$$m'(A_i) = m(A_i) \cdot (1+\epsilon) \tag{28}$$

Then, for the mass assignments of other focal elements in original BBA,

$$m'(A_j) = m(A_j) - \frac{m(A_j)}{1-m(A_i)} \cdot \epsilon \cdot m(A_i), \forall j \neq i \tag{29}$$

m' is the generated BBA with noise. It is easy to verify that

$$\sum\nolimits_{B\subseteq\Theta} m'(B) = 1 \tag{30}$$

It can be seen that the change of mass assignment for the primary focal element is the most significant when compared with those of other focal elements. The change of the mass assignment for primary focal element is redistributed to all the other focal elements according to the ratio among their corresponding mass assignments. BBAs are generated according to Algorithm 1 below [12].

For method I for adding noise, some restrictions should be adopted for the values of original BBA and the noise added to make sure that the noised BBA m' satisfies the defintion of BBA. The restriction are as shown in Eq. (31).

$$0 \leq (1+\epsilon) \cdot \max_{A}(m(A)) \leq 1, \forall A \in 2^{\Theta} \tag{31}$$

[2]For example, when $\Theta = \{\theta_1, \theta_2\}$ and $m(\{\theta_1\}) = 0.8, m(\{\theta_2\}) = 0.1, m(\Theta) = 0.1$, the primary focal element is $\{\theta_1\}$. When $\Theta = \{\theta_1, \theta_2\}$ and $m(\{\theta_1\}) = 0.45, m(\{\theta_2\}) = 0.45, m(\Theta) = 0.1$, the primary focal elements are $\{\theta_1\}$ and $\{\theta_2\}$.

Algorithm 1. Random generation of BBA

Input: Θ: Frame of discernment;
N_{max}: Maximum number of focal elements
Output: **Output**: m: BBA
Generate $\mathcal{P}(\Theta)$, which is the power set of Θ;
Generate a random permutation of $\mathcal{P}(\Theta) \rightarrow \mathcal{R}(\Theta)$;
Generate an integer between 1 and $N_{max} \rightarrow l$;
FOReach First k elements of $\mathcal{R}(\Theta)$ do
Generate a value within $[0,1] \rightarrow m_i$, $i = 1, ..., l$;
END
Normalize the vector $\mathbf{m} = [m_1, ..., m_l] \rightarrow \mathbf{m}'$;
$m(A_i) = m'_i$;

where m is the original BBA.

D. Method II for adding noise

Method II for adding noise is designed to evaluate the effect of creating new focal elements. Suppose that m is a BBA with a special structure defined on FOD Θ. The focal elements are some singletons $\{\theta_i\}$ and the total set Θ. First, find out a pair of singletons $\{\theta_i\}$ and $\{\theta_j\}$.

Second, create a new focal element $\{\theta_i, \theta_j\}$ with the mass value of ϵ, i.e., $m'(\{\theta_i, \theta_j\}) = \epsilon$.

Then, the mass values for focal elements $\{\theta_i\}$ and $\{\theta_j\}$ are regenerated as

$$\begin{cases} m'(\{\theta_i\}) = m(\{\theta_i\}) - \epsilon \cdot \frac{m(\{\theta_i\})}{m(\{\theta_i\})+m(\{\theta_j\})} \\ m'(\{\theta_j\}) = m(\{\theta_j\}) - \epsilon \cdot \frac{m(\{\theta_j\})}{m(\{\theta_i\})+m(\{\theta_j\})} \end{cases} \tag{32}$$

Obviously, one has $\sum_{B \subseteq \Theta} m'(B) = 1$.

The BBAs with special structure (with only some singletons and the total set focal elements) are generated according to Algorithm 2 below:

Algorithm 2. Random generation of BBA

Input: Θ: Frame of discernment;
n: Cardinality of Θ;
N_{max}: Maximum number of focal elements
Output: m: BBA
Generate $\mathcal{P}(\Theta)$, which is the power set of Θ;
Generate a random permutation of $\mathcal{P}(\Theta) \rightarrow \mathcal{R}(\Theta)$;
FOR $i = 1 : N_{max} - 1$
Generate an integers j between 1 and n;
Generate a focal element $F_i : \{\theta_j\}$;
END
Generate a focal element $F_{N_{max}} : \Theta$.
FOR $i = 1 : N_{max}$
Generate a value within $[0,1] \rightarrow m_i$;
END
Normalize the vector $m = [m_1, ..., m_{N_{max}}] \rightarrow m'$;
$m(F_i) = m'_i$;

For method II for adding noise, some restrictions should be adopted for the values of original BBA and the noise added to make sure that the noised BBA m' satisfies the defintion of BBA. According to Eq. (32), the restriction can be obtained as shown in Eq. (33). For all the available singleton focal element $\{\theta_i\}$ in original BBA,

$$0 \leq m(\{\theta_i\}) \cdot (1 - \frac{\epsilon}{\sum_{j, m(\{\theta_j\})>0} m(\{\theta_j\})}) \leq 1, \forall A \in 2^{\Theta} \tag{33}$$

where m is the original BBA.

E. A simple illustrative example

Here an illustrative example of single cycle calculation of the evaluation indices is provided by using Method I for adding noise. By referring to this illustrative example, evaluations by using Method II are easy to implement.

A BBA m defined on the FOD $\Theta = \{\theta_1, \theta_2, \theta_3\}$ is $m(\{\theta_1\}) = 0.6, m(\{\theta_2\}) = 0.3, m(\{\theta_1, \theta_2, \theta_3\}) = 0.1$.

Suppose that the noise sequence is
$\boldsymbol{\epsilon} = [-0.1, -0.05, -0.02, 0.02, 0.05, 0.1]$.

It can be seen that the restrictions in Eq. (31) are not violated.

According to the Step 1, we generate the sequence six noised BBA $\mathbf{m}' = [m_1, m_2, m_3, m_4, m_5, m_6]$ as follows:

$m_1(\{\theta_1\}) = 0.5, m_1(\{\theta_2\}) = 0.375, m_1(\{\theta_1, \theta_2, \theta_3\}) = 0.125;$
$m_2(\{\theta_1\}) = 0.55, m_2(\{\theta_2\}) = 0.3375, m_2(\{\theta_1, \theta_2, \theta_3\}) = 0.1125;$
$m_3(\{\theta_1\}) = 0.58, m_3(\{\theta_2\}) = 0.315, m_3(\{\theta_1, \theta_2, \theta_3\}) = 0.105;$
$m_4(\{\theta_1\}) = 0.62, m_4(\{\theta_2\}) = 0.285, m_4(\{\theta_1, \theta_2, \theta_3\}) = 0.095;$
$m_5(\{\theta_1\}) = 0.65, m_5(\{\theta_2\}) = 0.2625, m_5(\{\theta_1, \theta_2, \theta_3\}) = 0.0875;$
$m_6(\{\theta_1\}) = 0.7, m_6(\{\theta_2\}) = 0.225, m_6(\{\theta_1, \theta_2, \theta_3\}) = 0.075.$

Here we use Dempster's rule of combination. Then, according to the Step 2, the original combination sequence $\mathbf{m}_c = [m_c^1, m_c^2, ..., m_c^6]$ is as

$\forall i = 1, ..., 6.$
$m_c^i(\{\theta_1\}) = 0.75, m_c^i(\{\theta_2\}) = 0.2344, m_c^i(\{\theta_1, \theta_2\, \theta_3\}) = 0.0156.$

Then according to the Step 3, the sequence of combination results by combining BBAs with noise and the original BBAs $\mathbf{m}_{cn} = [m_{cn}^1, m_{cn}^2, ..., m_{cn}^6]$ is as

$m_{cn}^1(\{\theta_1\}) = 0.6627, m_{cn}^1(\{\theta_2\}) = 0.3162,$
$m_{cn}^1(\{\theta_1, \theta_2, \theta_3\}) = 0.0211;$
$m_{cn}^2(\{\theta_1\}) = 0.7076, m_{cn}^2(\{\theta_2\}) = 0.2741,$
$m_{cn}^2(\{\theta_1, \theta_2, \theta_3\}) = 0.0183;$
$m_{cn}^3(\{\theta_1\}) = 0.7334, m_{cn}^3(\{\theta_2\}) = 0.2500,$
$m_{cn}^3(\{\theta_1, \theta_2, \theta_3\}) = 0.0167;$
$m_{cn}^4(\{\theta_1\}) = 0.7662, m_{cn}^4(\{\theta_2\}) = 0.2192,$
$m_{cn}^4(\{\theta_1, \theta_2, \theta_3\}) = 0.0146;$
$m_{cn}^5(\{\theta_1\}) = 0.7895, m_{cn}^5(\{\theta_2\}) = 0.1973,$
$m_{cn}^5(\{\theta_1, \theta_2, \theta_3\}) = 0.0132;$
$m_{cn}^6(\{\theta_1\}) = 0.8257, m_{cn}^6(\{\theta_2\}) = 0.1634,$
$m_{cn}^6(\{\theta_1, \theta_2, \theta_3\}) = 0.0109.$

For the noise sequence,

$$\|\boldsymbol{\epsilon}\|^2 = 0.0258$$

$$Var(\boldsymbol{\epsilon}) = 0.0043$$

According to the Step 4, the value of MSE is

$$\mathrm{MSE}_{\mathrm{BBA}}(\mathbf{m}_{cn}) = 0.00270$$

$$\mathrm{MSE}'_{\mathrm{BBA}}(\mathbf{m}_{cn}) = 0.00270/0.0258 = 0.104752$$

According to the Step 5, the value of variance is

$$\mathrm{Var}_{\mathrm{BBA}}(\mathbf{m}_{cn}) = 0.002697$$

$$\mathrm{Var}'_{\mathrm{BBA}}(\mathbf{m}_{cn}) = 0.002697/0.0043 = 0.6272$$

In the final, according to the Step 5, the value of bias is

$$\mathrm{Bias}_{\mathrm{BBA}}(\mathbf{m}_{cn}) = 0.002406$$

$$\mathrm{Bias}'_{\mathrm{BBA}}(\mathbf{m}_{cn}) = 0.002406/\sqrt{0.0258} = 0.014978$$

The above is the illustration of one-cycle procedure. One can use other combination rules to do these steps. Randomly generate BBAs and repeat all the steps, then we can obtain the final statistical evaluation results.

VI. Simulations

A. Simulation I: using Method I for adding noise

In our simulations, the cardinality of the FOD is 3. In random generation of BBAs, the number of focal elements has been set to 5. The length of the noise sequence is 50 (the noise value starts from -0.1, with an increasing step of 0.004, up to 0.1. Of course, the zero value for noise is not considered because it corresponds to noiseless case.). In each simulation cycle, seven combination rules including Dempster's rules and other alternatives aforementioned in Section III are used, respectively. We have repeated the Monte Carlo simulation with 100 runs. In random generation of original BBAs, the restrictions in Eq. (31) are not violated. The statistical results are listed in Tables I-III. The ranks of the relative MSE, relative variance and relative bias are obtained based on the descending order.

It should be noted that when using RCR in our simulation, the weights are generate as follows.

$$\begin{cases} \alpha(K) = \frac{K}{1-K+K^2} \\ \beta(K) = \frac{1-K}{1-K+K^2} \end{cases} \tag{34}$$

TABLE I. Comparisons in terms of MSE

Combination Rules	$\mathrm{MSE}'_{\mathrm{BBA}}$	Rank
Dempster's rule	0.0010758	1
Yager's rule	0.0005743	5
Disjunctive rule	0.0004298	7
D&P rule	0.0006247	4
RCR	0.0007152	3
PCR5	0.0010505	2
Mean rule	0.0005711	6

TABLE II. Comparisons in terms of Variance

Combination Rules	$\mathrm{Var}'_{\mathrm{BBA}}$	Rank
Dempster's rule	1.7260	1
Yager's rule	0.9568	6
Disjunctive rule	0.7469	7
D&P rule	1.0226	4
RCR	1.2788	3
PCR5	1.6801	2
Mean rule	1	5

As we can see in Tables I - III, Dempster's rule are the most sensitive to the mass change according to the criterion of the relative bias, and it also has highest degree of divergence according to the criterion of relative variance. Mean rule is the most insensitive to the mass change according to the criteria

TABLE III. Comparisons in terms of Bias

Combination Rules	$\mathrm{Bias}'_{\mathrm{BBA}}$	Rank
Dempster's rule	$0.81132*10^{-7}$	1
Yager's rule	$0.59228*10^{-7}$	2
Disjunctive rule	$0.41949*10^{-7}$	4
D&P rule	$0.49075*10^{-7}$	3
RCR	$0.39592*10^{-7}$	5
PCR5	$0.38066*10^{-7}$	6
Mean rule	0	7

of relative bias, and it is always a rule with smaller divergence according to the criterion of the relative variance. Yager's rule is always more sensitive to the mass change and is always not so divergent.PCR5 rule is not so sensitive to the mass change according to the criterion of Bias (rank 6), and it is not so divergent according to the criterion of the relative variance. The Robust combination rule (RCR), Dubois & Prade's rule (D&P rule) are always moderate to the mass change in terms of sensitivity and in terms of divergence. So, PCR5 and RCR are more moderate rules; thus, they are relatively good choices for practical use.

B. Simulation II: using Method II for adding noise

In our simulations, the cardinality of the FOD is 3. In generation of BBAs, the total set Θ is used as a focal element and the number of singleton focal elements has been set to 2. The length of the noise sequence is 50 (the noise value starts at 0.002 with an increasing step of 0.002, up to 0.1.) In each simulation cycle, seven combination rules including Dempster's rules and other alternatives aforementioned in Section III are used, respectively. We have repeated the Monte Carlo simulation with 100 runs. In random generation of original BBAs, the restrictions in Eq. (33) are not violated. The statistical results are listed in Tables IV-VI. The ranks of the relative MSE, relative variance and relative bias are obtained based on the descending order.

The derivation of weights of RCR has been done in the same manner as for the Simulation I.

TABLE IV. Comparisons in terms of MSE

Combination Rules	$\mathrm{MSE}'_{\mathrm{BBA}}$	Rank
Dempster's rule	0.0050	3
Yager's rule	0.0038	5
Disjunctive rule	0.0033	6
D&P rule	0.0014	7
RCR	0.0078	1
PCR5	0.0043	4
Mean rule	0.0056	2

TABLE V. Comparisons in terms of Variance

Combination Rules	$\mathrm{Var}'_{\mathrm{BBA}}$	Rank
Dempster's rule	0.9218	3
Yager's rule	0.7019	5
Disjunctive rule	0.5434	6
D&P rule	0.2714	7
RCR	1.3528	1
PCR5	0.8014	4
Mean rule	1	2

As we can see in Tables IV - VI, RCR is the most sensitive to the change of focal elements according to the criterion of the relative bias, and it also has highest degree

TABLE VI. COMPARISONS IN TERMS OF BIAS

Combination Rules	$\text{Bias}'_{\text{BBA}}$	Rank
Dempster's rule	0.0016	3
Yager's rule	0.0012	5
Disjunctive rule	0.0011	6
D&P rule	0.0004	7
RCR	0.0025	1
PCR5	0.0013	4
Mean rule	0.0018	2

of divergence according to the criterion of relative variance. Dubois & Prade's rule (D&P rule) is the most insensitive rule according to the criteria of relative bias, and it is always a rule with smaller divergence according to the criterion of the relative variance. Yager's rule is always insensitive and is always not so divergent. Mean rule is sensitive to the change of focal element according to the criterion of Bias (rank 2), and it is divergent according to the criterion of the relative variance. Dempster's rule is not so sensitive to the change of focal element. The PCR5 and Yager's rules are always moderate to the change of focal elements in terms of sensitivity and in terms of divergence.

According to simulations results, we see that the different methods of adding noises impact differently the results of the comparative evaluations. However, we have shown that no matter the method adopted (by keeping the original core of the BBA, or modifying it slightly), PCR5 provides quite robust results for combining two BBA's and thus offers practical interests from this standpoint.

VII. CONCLUSION

In this paper we have proposed a group of statistical criteria for evaluating the sensitivity of different combination rules with respect to the noise perturbations. The design is based on the classical measures of performance like MSE, variance, and bias encountered in the estimation theory. We don't rank the rules according to their a priori "good expected" properties. Moderate relative bias values are preferred, which means the balance or trade-off between the robustness and the sensitivity. Small relative variance values are preferred, which represent the high cohesion of a given combination rule. Seven widely used evidence combination rules were evaluated using the new proposed evaluation criteria. PCR5 is a moderate rule which is good for the practical use for combining two BBAs. For combining more than two BBAs, we expect that PCR6 will be a good choice, but we need to make more investigations in future to evaluate precisely its performances.

In this work, we have added some noises to BBAs mainly by modifying the mass assignments of the primary focal element and by creating new focal elements. In our future work, we will try to use other methods to add noise to BBAs, e.g., eliminating some of original focal elements. In our Monte-Carlo simulations, there is no pre-settings of mass assignments for the BBAs. In this paper, in each cycle we only generate one BBA, based on which, we generate a sequence of BBAs by adding small noise. The BBAs to be combined are the original BBAs and the BBAs with small noise. In our future work, we will try to generate two BBA sequences and add noise to them, respectively, where we can use some special BBAs in the evaluation procedure, e.g., BBAs to be combined are high conflicting. Then we can do more specific performance evaluations on the combination rules. In this paper, we did only focus on the property of sensitivity and divergence. The evaluation criteria of other aspects of evidence combination are also required for evaluating and designing new combination rules, which will be investigated in future research works and forthcoming publications.

ACKNOWLEDGMENT

This work was supported by the Grant for State Key Program for Basic Research of China (973) No. 2013CB329405, National Natural Science Foundation (No. 61104214, No. 61203222), Specialized Research Fund for the Doctoral Program of Higher Education (20120201120036), Shaanxi Province Science and Technology Research and Development Program (2013KJXX-46), and Fundamental Research Funds for the Central Universities.

REFERENCES

[1] G. Shafer. *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press, 1976.

[2] L. A. Zadeh. "A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination," AI magazine, 1986, vol. 7, no. 2, p. 85-90.

[3] J. Dezert, P. Wang, A. Tchamova. "On the validity of Dempster-Shafer theory," In: Proceedings of the 15th International Conference on Information Fusion, Singapore, July, 2012, p. 655-660.

[4] R. R. Yager. "On the Dempster-Shafer framework and new combination rules," Information Science, 1987, vol. 41, no. 2, p. 93-137.

[5] M. C. Florea, A.-L. Jousselme, E. Bossé, D. Grenier. "Robust combination rules for evidence theory," Information Fusion, 2009, vol. 10, no. 2, p. 183-197.

[6] P. Smets. "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem," International Journal of Approximate reasoning, 1993, vol. 9, no. 1, p. 1-35.

[7] D. Dubois, H. Prade. "Representation and combination of uncertainty with belief functions and possibility measures," Computational Intelligence, 1988, vol. 4, no. 3, p. 244-264.

[8] F. Smarandache, J. Dezert. *Applications and Advances of DSmT for Information Fusion (Vol3)*. Rehoboth: American Research Press, 2009.

[9] C. K. Murphy. "Combining belief functions when evidence conflicts," Decision Support Systems, 2000, vol. 29, no. 1, p. 1-9.

[10] R. Haenni. "Are alternatives to Dempster's rule of combination real alternative?: Comments on 'About the belief function combination and the conflict management problem' ," Information Fusion, 2002, vol. 3, no. 4, p. 237-239.

[11] Y. Bar-Shalom, X. Rong Li, T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. USA, New York: John Wiley & Sons, 2001.

[12] A.-L. Jousselme, P. Maupin. "Distances in evidence theory: Comprehensive survey and generalizations," International Journal of Approximate Reasoning, 2012, vol. 53, no. 2, p. 118-145.

[13] A.-L. Jousselme, D. Grenier, E. Bossé. "A new distance between two bodies of evidence," Information Fusion, 2001, vol. 2, no. 2, p. 91-101.

[14] D. Q. Han, Y. Deng, C. Z. Han, Y. Yang. "Some notes on betting commitment distance in evidence theory," Science China - Information Sciences, 2012, vol. 55, no. 3, p. 558-565.

[15] M. Bouchard, A.-L. Jousselme, P.-E. Doré. "A proof for the positive definiteness of the Jaccard index matrix," International Journal of Approximate Reasoning, 2013, vol. 54, no. 5, p. 615-626.

[16] F. Smarandache, J. Dezert. "On the consistency of PCR6 with the averaging rule and its application to probability estimation," In Proc. of Fusion 2013 Int. Conference on Information Fusion, Istanbul, Turkey, July 9-12, 2013, p. 1127-1134.

[17] F. Smarandache, J. Dezert. *Applications and Advances of DSmT for Information Fusion (Vol2)*. Rehoboth: American Research Press, 2006.

A New Self-Adaptive Fusion Algorithm Based on DST and DSmT

Xiao-Hong Yu
Qing-Jun Zhou
Yan-Li Li
Jin An
Zhi-Cheng Liu

Originally published as Yu X.H., Zhou Q.-J., Li Y.-L., An J., Liu Z.-C., *A new self-adaptive fusion algorithm based on DST and DSmT,* in Proc. of 17th Int Conf on Information Fusion, Salamanca, Spain, July 7-10, 2014, and reprinted with permission.

***Abstract*—A new self-adaptive fusion algorithm based on DST and DSmT is proposed. In the new algorithm, part of the conflicting information is normalized according to DST, while the other part is processed by DSmT. A controlling factor is used to control the quantity of information dealt by the two different methods adaptively, which is a new method avoiding setting for the threshold of conflict. The simulation results indicate that the new self-adaptive fusion algorithm based on DST and DSmT can deal with any conflicting situation with a good performance of convergence.**

***Keywords*—DST; DSmT; controlling factor; information fusion**

I. INTRODUCTION

Due to the complexity of modern battlefield, target identification is becoming more and more complex. It is difficult to give an accurate and credible identifying result only by one sensor. Therefore, target identification based on multi-source information is becoming a hot topic. Dempster-Shafer theory (DST) is an efficient method for uncertainty consequence ([1]). It is widely used in the domain of synthesize identification. However, DST can't give efficient fusion results when information from different sources becomes highly conflict. Many improvement are proposed, such as Yager's rule of combination ([2]), Murphy's rule of combination ([3]), Dengyong's rule of combination ([4]) etc. Dezert presented the Dezert-Smarandache theory (DSmT) ([5]), which can be considered as an extension of the classical DST. DSmT performs well in dealing with the fusion of uncertain, highly conflicting and imprecise sources. It can solve not only the static problems but also the complex dynamic fusion problems.

DST and DSmT have their own advantages and disadvantages for fusing the multi-source information. The advantages of DST mainly occur in the case of low degree of conflict, whereas it may give a bad fusion result which is absolutely contrary to the fact while the sources are in high degree of conflict. DSmT is more efficient in combining highly conflicting sources, but it offers convergence toward certainty slowly especially in low degree of conflict. So a new self-adaptive fusion algorithm is put forward in this paper.

II. REVIEW OF THE THEORY OF EVIDENCE

A. DST

DST was firstly proposed by Dempster in 1967 and extended by Shafer. The main idea will be reviewed as follows. Let $\Theta = \{\theta_1, \cdots, \theta_n\}$ be the frame of discernment of the fusion problem and all elements of Θ are exclusive. A basic belief assignment (BBA) $m: 2^\Theta \to [0,1]$ is defined as

$$\begin{cases} \sum_{A \in 2^\Theta} m(A) = 1 \\ m(\phi) = 0 \end{cases} \tag{1}$$

where 2^Θ is the power set of Θ and it includes all its subsets .

For two independent bodies of evidence whose BBAs are denoted by m_1 and m_2 respectively, the BBA of the combination of the two bodies is given by the following rule

$$m_{12}(X) = \begin{cases} \dfrac{\sum_{A,B \in 2^\Theta, A \cap B = X} m_1(A) m_2(B)}{1-k}, \forall X \in \Theta, X \neq \phi \\ 0, X = \phi \end{cases} \tag{2}$$

where $k = \sum_{A \cap B = \phi} m_1(A) m_2(B)$ reflects the conflict degree of the two sources.

B. DSmT

DSmT, an extension of DST, was developed by Dezert and Smarandanche ([5]). DSmT differs from DST at that the elements of Θ could be overlapped. For simplicity, use D^Θ (Hyper-power set) to denote the set of all compositions built from elements of Θ with $\bigcup$ and $\bigcap$ operators. The generalized basic belief assignment (GBBA) $m: D^\Theta \to [0,1]$ is defined as

$$\begin{cases} \sum_{A \in D^\Theta} m(A) = 1 \\ m(\phi) = 0 \end{cases} \tag{3}$$

Similarly, the classical combination rule for two independent bodies of evidence, whose GBBAs are m_1 and m_2 respectively, is given by

$$m_{M^f(\theta)}(X)=\begin{cases}\sum_{A,B\in D^\Theta, A\cap B=X} m_1(A)m_2(B), \forall X\in D^\Theta, X\neq\phi \\ 0, X=\phi\end{cases} \quad (4)$$

It is remarked that DSmT keeps the conflicting information and doesn't need normalization.

The biggest difference between DST and DSmT can be intuitively described as follows. Let $\Theta=\{\theta_1,\theta_2\}$ be the frame of discernment (without any extra conditions). The DST deals with BBA $m(\cdot)\in[0,1]$ such that $m(\theta_1)+m(\theta_2)+m(\theta_1\cup\theta_2)=1$, while the DSmT deals with GBBA $m(\cdot)\in[0,1]$ such that $m(\theta_1)+m(\theta_2)+m(\theta_1\cup\theta_2)+m(\theta_1\cap\theta_2)=1$.

C. Self-adaptive fusion algorithm based on DST and DSmT

Although DST behaves well while fusing sources are in low degree of conflict, it may give a fusion result that absolutely contrary to the fact while two sources are in high degree of conflict. Luckily, DSmT is capable of fusing the highly conflicting sources. A natural idea comes out that a self-adaptive fusion algorithm based on DST and DSmT could be used to obtain better performance in a way of simple combination, that is, if the degree of conflict is less than a given threshold, Dempster-Shafer combination rule can be used, otherwise DSm combination rule will be selected.

The limitation of applying the above idea in practice is that a threshold of conflict should be set in advance. It is very difficult to determine a suitable value for the conflicting threshold since different systems have different degrees of conflict, and an experiential value is usually used instead the true one by experimenting time after time. The risk lies that once the threshold is not suitable, fusion results will be bad.

To solve the above problem, this paper proposed a new self-adaptive algorithm based on DST and DSmT, which avoids setting for the conflicting threshold in advance.

III. A New Self-adaptive Fusion Algorithm

Note that the key difficulty lies in how to process the conflicting information. Before introducing the new self-adaptive fusion algorithm, we will review the existing methods and explain why we adopt such an algorithm.

According to DST, conflicting pieces of information between two bodies of evidence are eliminated by making normalization. Yager ([2]) distributed the conflicting information to the union of all elements, and viewed the conflict as unreliable and ignored it. Smets ([8]) distributed the conflicting information to empty set. He pointed out that all the sources of evidence are reliable but the frame of discernment is not complete and the actual result may lie out of the frame. The limitation of DST method is that it cannot deal with the cases of high degree conflicts.

On the contrary, DSmT keeps the conflicting information useful and redistributes the conflicting information according to some principles while fusing. It is capable of coping with the cases of high degree conflicts, but it offers slow convergence for the result.

To absorb the advantages of DST and DSmT, the new algorithm will treat the conflicting information in a combination way, that is, part of the conflicting information will be distributed to the nonempty set averagely and the other will be redistributed by other principles. A controlling factor will be proposed to decide the mass of conflicting information to be normalized or not.

A. Evaluation of conflict between bodies of evidence

Definition 1 (Conflict)[6]. A conflict between two beliefs in DS theory can be interpreted qualitatively as one source strongly supports one hypothesis and the other strongly supports another hypothesis, and the two hypotheses are not compatible.

It is well known that, the key problem of designing self-adaptive fusion algorithm based on DST and DSmT is how to compute the conflict between two bodies of evidence. Next we will show that the existing measures, including the conflict coefficient used in DST and DSmT and the degree of similarity, are not suitable to act as a eligible measure for conflicting information, although the former has long been taken as a fact in the Dempster–Shafer theory community. We also propose a new measure to fill in this gap.

Example 1. Let Θ be a frame of discernment with n hypotheses $\{\theta_1,\cdots,\theta_n\}$. Assume m_1 and m_2 are two BBAs offered by two distinct sources which are defined as

$$m_1(\theta_i)=1/n,\quad m_2(\theta_i)=1/n,\quad i=1,2,\cdots,n$$

Obviously, the two BBAs are totally consistent with each other. So there shouldn't exist conflict. Firstly, compute the conflict coefficient and get $k=1-1/n$, where n is the number of hypotheses in frame of discernment. Secondly, It can be found out that along with the increase of n, k will increase and approach to 1, as shown in the relationship between k and n is shown as Fig.1. If k is taken as a measurement of the degree of conflict, the two bodies of evidence are in high degree of conflict when n is larger than 5. It is surely contrary to the fact.

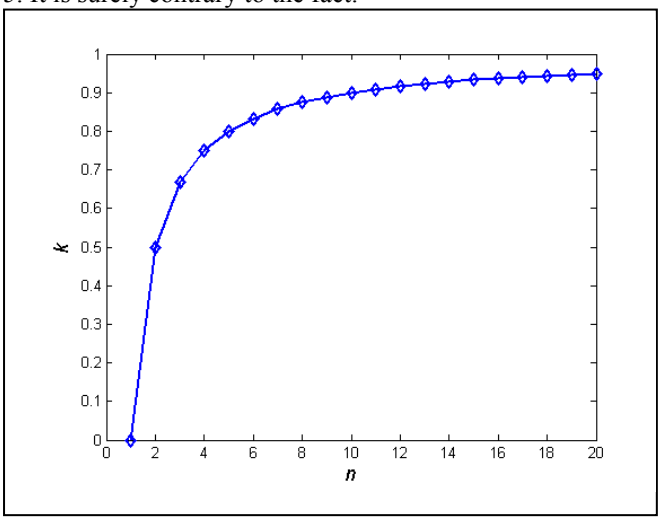


Fig.1. Relationship between k and n in example 1

Example 2. Let Θ be a frame of discernment with two hypotheses $\{\theta_1, \theta_2\}$. The BBAs offered by two sources are defined as

$$m_1(\theta_1) = p,\ m_1(\theta_2) = 1-p$$
$$m_2(\theta_1) = 1-p,\ m_2(\theta_2) = p$$

where $p \in [0,1]$. It is obvious that the two bodies of evidence are highly contradicted with each other. According to DST, the conflict coefficient can be calculated as $k = p^2 + (1-p)^2$. Fig 2 shows the relationship between k and p. As p changes from 0 to 1, k will decrease from 1 to 0.5 and then increase to 1 again, as shown in Fig 2. Note that k is low than 1 at most time especially while p is around 0.5. It can't reflect the conflict between two sources of evidence rightly.

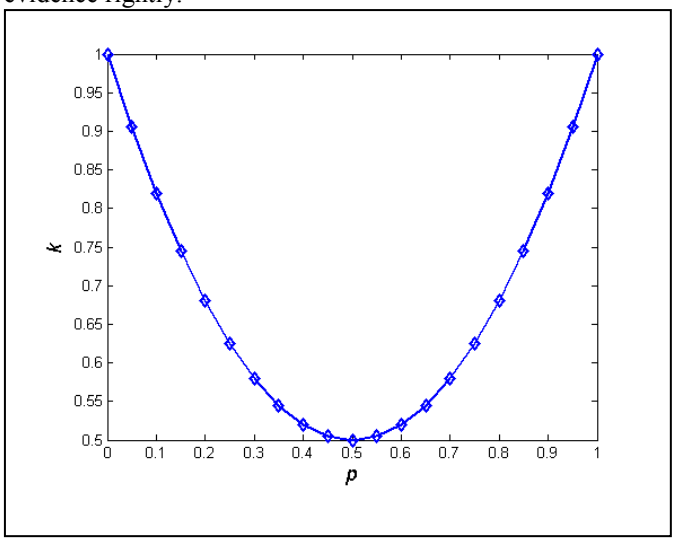


Fig.2. Relationship between k and n in example 2

From Example 1 and 2, we can see that the conflict coefficient k can't be used as a suitable measure of conflict.

To overcome the above shortage, some other evaluating methods of conflict are put forward, such as the degree of similarity. Since the similarity of two bodies of evidence can also reflect their conflict, it can be used as a candidate to reveal the conflict between two sources of evidence. Generally speaking, the larger the degree of similarity is, the smaller the conflict is.

Usually, the degree of similarity can be calculated by the distance between bodies of evidence. There are some kinds of distance being used in information fusion, such as the famous the Euclidean distance ([7]) proposed by Cuzzolin, the Bhattacharyya distance ([8]) given by Ristic and Smets. Both of them are defined in the frame of Dempster-Shafer theory. Nevertheless, neither of them can reflect the similarity of the subset of frame Θ. Besides, Tessem turned the belief function into the probability function by pignistic transformation and evaluated the distance between bodies of evidence in the level of pignistic ([9]). However, the distance defined in this way doesn't accord with the distance theory. No valuable distance can be used in practice until the coming out of Jousselme distance ([10]), which is the most widely used distance of evidence at present. The definition of Jousselme distance is given as follows.

Definition 2 (Jousselme distance). Let Θ be a frame of discernment with n hypotheses. The BBAs offered by two sensors are denoted as m_1 and m_2. Distance between them can be defined as

$$Dis(m_1, m_2) = \sqrt{\frac{1}{2}(m_1 - m_2)^T D (m_1 - m_2)} \qquad (5)$$

where $D = (D_{ij}) = \left(\frac{|A_i \cap B_j|}{|A_i \cup B_j|} \right)$ is a $2^n \times 2^n$-dimensional matrix, and $|A|$ denotes the number of elements of A. It reflects the degree of similarity of the evidence. Formula (5) can be rewritten as

$$Dis(m_1, m_2) = \sqrt{\frac{1}{2}(\| m_1 \|^2 + \| m_2 \|^2 - 2 < m_1, m_2 >)}$$

where $\| m_i \|^2 = < m_i, m_i >$, $i = 1, 2$, and

$$< m_1, m_2 > = \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} m_1(A_i) m_2(B_j) \frac{|A_i \cap B_j|}{|A_i \cup B_j|}$$

is the product of two vectors.

Note that in the frame of discernment in DSmT, hypotheses could be overlapped potentially. Then a generalized Jousselme distance is defined as follows.

Definition 3 (generalized Jousselme distance). Let Θ be a fame of discernment in DSmT with n hypothesis. The two GBBAs offered by sensors are denoted as m_1 and m_2. Distance between them is defined as

$$Dis(m_1, m_2) = \sqrt{\frac{1}{2}(m_1 - m_2)^T D (m_1 - m_2)} \qquad (6)$$

where $D = (D_{ij}) = \left(\frac{|A_i \cap B_j|}{|A_i \cup B_j|} \right)$ is a $N \times N$-dimensional matrix, N is the number of elements in the power set of Θ, and $|A|$ denotes the DSm cardinality of A. Similarly, distance can be transformed as

$$Dis(m_1, m_2) = \sqrt{\frac{1}{2}(\| m_1 \|^2 + \| m_2 \|^2 - 2 < m_1, m_2 >)},$$

where $\| m_i \|^2 = < m_i, m_i >$, $i = 1, 2$, and

$$< m_1, m_2 > = \sum_{i=1}^{N} \sum_{j=1}^{N} m_1(A_i) m_2(B_j) \frac{|A_i \cap B_j|}{|A_i \cup B_j|},$$

$\forall A \in D^{\Theta}$.

It is easy to see that $Dis(m_1, m_2) \in [0,1]$, and the degree of similarity can be defined as

$$Sim(m_1, m_2) = 1 - Dis(m_1, m_2) \quad (7)$$

Obviously, it has $Sim(m_1, m_2) \in [0,1]$.

In example 1, the Jousselme distance is always 0 no matter what the value of n be. In other words, the similarity between two bodies of evidence is always 1, it accords with the fact. In example 2, the Jousselme distance is computed as $Dis(m_1, m_2) = |1-2p|$. As p changes from 0 to 1, the Jousselme distance decreases from 1 to 0 and then increases to 1 again. It is also reasonable in intuition. However, one cannot determine the value of conflict between two bodies of evidence correctly just by the similarity. We will show this by the following Example 3.

Example 3. Let Θ be a frame of discernment with two hypotheses $\{\theta_1, \theta_2\}$. The two BBAs offered by sensors are defined as

$$m_1(\theta_1) = 0.8,\ m_1(\theta_1 \bigcup \theta_2) = 0.2$$
$$m_2(\theta_1 \bigcup \theta_2) = 1$$

Simple calculation will yield $Dis(m_1, m_2) = 0.5657$, and $Sim(m_1, m_2) = 0.4343$. If the conflict is estimated by the degree of similarity, the two sources of evidence are in conflict. However, it is obvious that the second body of evidence is totally unknown. Thus one can't assert that they are in conflict. In other words, it is not credible to determine the degree of conflict only by on the degree of similarity.

In summary, neither conflict coefficient nor degree of similarity can be used as the quantitative measure of conflict alone. A natural idea comes out that one may make a judgment objectively by considering the two factors synthetically. Actually, many researchers followed this way. For example, Jiang ([12]) took the average of Joussleme distance and conflict coefficient as the new measure for conflict. Liu ([6]) made the dualistic array by conflict coefficient and the distance between betting commitments to analysis the conflict under different situations. Liu ([13]) used the geometric mean of conflict coefficient and distance between betting commitments as the measure of the conflict.

This paper also adopts the above idea. One will see in the following text that, in our new combination model there are two places where the degrees of conflict need to be estimated. On one side, the classical conflict coefficient is taken to measure the value of conflict. On the other side, the similarity between bodies of evidence is used as a controlling factor to distribute the conflicting information. See next for details.

B. New combination rule

Let $\Theta = \{\theta_1, \cdots, \theta_n\}$ be a discernment frame with n hypotheses. The hypotheses of the frame could be non-exclusive. D^Θ is the hyper-power set. The generalized basic belief assignment is defined as $m: D^\Theta \to [0,1]$ where

$$\begin{cases} \sum_{A \in D^\Theta} m(A) = 1 \\ m(\phi) = 0 \end{cases} \quad (8)$$

Let m_1 and m_2 be the GBBAs of two sensors which are independent with each other. The new combination rule can be defined as

$$m_{12}(X) = \frac{\sum_{\substack{A,B \in D^\Theta \\ A \bigcap B = X}} m_1(A) m_2(B) + P(X)}{1 - k'}, \forall X \in D^\Theta, X \neq \phi \quad (9)$$

where $k = \sum_{A,B \in D^\Theta, A \bigcap B = \phi} m_1(A) m_2(B)$ reflects the mass of conflict, $k' = \sigma k$ is the conflict that will be distributed to the all hypotheses averagely by normalizing. The rest $(1-\sigma)k$ will be redistributed by other rules. Here σ is a controlling factor. Denote the conflicting information that be distributed to hypothesis X by $P(X)$, and thus $\sum_{X \in D^\Theta} P(X) = (1-\sigma)k$.

It can be seen from formula (9) that, as σ changes, the fusion results will be different. When $\sigma = 0$, all conflict information will be kept and distributed to the hypotheses which bring on the conflict, and then the new algorithm will degenerate to DSm combination rule. When $\sigma = 1$, all conflict information will be distributed to all hypotheses averagely, and then the new algorithm will degenerate to DS combination rule. When $\sigma \in (0,1)$, part of the conflict will be kept as useful and the rest will be distributed averagely, then the new algorithm is the synthesis of DSm and DS combination rules.

The new algorithm uses the degree of similarity as the controlling factor to adjust the fusion result adaptively. Let s be the number of sources. If $s = 2$, then $\sigma = Sim(m_1, m_2)$; if $s > 2$, then $\sigma = \min\{Sim(m_i, m_j) \mid i, j = 1, \cdots, s\}$.

While there are more than two sources to be fused, the new combination rule can be defined as

$$m(X) = \frac{\sum_{\bigcap A_i = X} \prod_{1 \le j \le s} m_j(A_i) + P(X)}{1 - k'}, \forall X \in D^\Theta, X \neq \phi \quad (10)$$

where $k' = \sigma k$, $k = \sum_{\bigcap A_i = \phi} \prod_{1 \le j \le s} m_j(A_i)$.

C. Conflict distribution rule

To copy with the complex constraints in real systems, Dezert proposed the hybrid DSm combination rule, which works properly even if in high degree of conflict. However, due to the big number of element in D^Θ, it cannot offer quick convergence and cost too much time for calculation. To achieve a better performance, some new distribution rules based on the DSm rule are put forward and are classified as PCR1~PCR6 according to distribution rules ([11]). It is remarked that PCR5 is thought to be the most precise in distribution and its combination rule is defined as

$$m(X)=\sum_{A\cap B=X} m_1(A)m_2(B) + \sum_{\substack{Y\in D^\Theta \\ X\cap Y=\phi}}\left[\frac{m_1^2(X)m_2(Y)}{m_1(X)+m_2(Y)}+\frac{m_2^2(X)m_1(Y)}{m_2(X)+m_1(Y)}\right] \quad (11)$$

where $X\in D^\Theta, X\neq\phi$.

According to PCR5, the conflicting information caused by X and Y will be distributed to themselves without considering of their union. In fact, the conflicting information is decompounded to two parts as $m_1(X)m_2(Y)$ and $m_2(X)m_1(Y)$, which will be distributed separately. Hypotheses X and Y will get the conflicting information in proportion to their basic belief assignments. Due to the high performance, PCR5 is widely used in real systems. As an upgraded version of PCR5, PCR6 is the latest rule which is used to fuse more bodies of evidence and has also been applied in real systems.

Our algorithm also adopts the PCR5 for conflict distribution, and then the item $P(X)$ in formula (9) could be rewritten as:

$$P(X)=(1-\sigma)\sum_{\substack{Y\in D^\Theta \\ X\cap Y=\phi}}\left[\frac{m_1^2(X)m_2(Y)}{m_1(X)+m_2(Y)}+\frac{m_2^2(X)m_1(Y)}{m_2(X)+m_1(Y)}\right] \quad (12)$$

If there are more than two sources, they can be combined one by one according to PCR5. In addition, they can also be combined according to PCR6, and the item $P(X)$ in formula (9) could be rewritten as

$$P(X)=(1-\sigma)\left\{\sum_{i=1}^{s} m_i^2(X)\sum_{\bigcap_{k=1}^{s-1} Y_{\sigma_i(k)}\cap X=\phi}\left[\frac{\prod_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)})}{m_i(X)+\sum_{j=1}^{s-1} m_{\sigma_i(j)}(Y_{\sigma_i(j)})}\right]\right\} \quad (13)$$

where $\sigma_i(j)=\begin{cases} j, j<i \\ j+1, j\geq i\end{cases}$.

D. Realization of the new algorithm

Let $\Theta=\{\theta_1,\cdots,\theta_n\}$ be a frame of discernment with n hypotheses. $m_1(A_i)$ and $m_2(B_j)$ are the basic belief functions of two sources which are independent with each other, where $\sum_{A_i\in D^\Theta} m_1(A_i)=1$, $\sum_{B_j\in D^\Theta} m_2(B_j)=1$. The new self-adaptive fusion algorithm based on DST and DSmT can be described as following four steps. Here we use $m(X)$ to denote the fusion result.

Step1. $\forall X\in D^\Theta, X\neq\phi$, set $m(X)=0$ and $k=0$. Compute $Dis(m_1,m_2)$, $Sim(m_1,m_2)$, and $\sigma=Sim(m_1,m_2)$.

Step2. $\forall i,j=1,2,\cdots$, compute $m_1(A_i)m_2(B_j)$. If $A_i\cap B_j\neq\phi$, then renew $m(A_i\cap B_j)$ with $m(A_i\cap B_j)+m_1(A_i)m_2(B_j)$, else renew k with $k+m_1(A_i)m_2(B_j)$, renew $m(A_i)$ with $m(A_i)+(1-\sigma)m_1^2(A_i)m_2(B_j)/(m_1(A_i)+m_2(B_j))$, and renew $m(B_j)$ with $m(B_j)+(1-\sigma)m_1(A_i)m_2^2(B_j)/(m_1(A_i)+m_2(B_j))$;

Step3. If all of $m_1(A_i)m_2(B_j)$ have been computed, go to step4; otherwise, go to step 2;

Step4. $\forall X\in D^\Theta$, if $X=\phi$, $m(X)=0$; otherwise, $m(X)=m(X)/(1-\sigma k)$.

IV. NUMERICAL EXAMPLE

Suppose five sensors are used to detect targets which are independent with each other. Let $\Theta=\{A,B,C\}$ be the frame of discernment. Elements in Θ are exclusive.

Example 1 (no conflicting evidence): The basic belief assignments offered by five sensors are given as follows.

Evidence 1: $m_1(A)=0.5, m_1(B)=0.2, m_1(C)=0.3$;

Evidence 2: $m_2(A)=0.6, m_2(B)=0.2, m_2(C)=0.2$;

Evidence 3: $m_3(A)=0.55, m_3(B)=0.1, m_3(C)=0.35$;

Evidence 4: $m_4(A)=0.55, m_4(B)=0.1, m_4(C)=0.35$;

Evidence 5: $m_5(A)=0.55, m_5(B)=0.1, m_5(C)=0.35$;

It is easy to see that all bodies of evidence support the identity A and they are in low degree of conflict. Fusion results offered by different combination rules are given in table 1.

Example 2 (one conflicting body of evidence): The basic belief assignments offered by five sensors are given as follows.

Evidence 1: $m_1(A)=0.5, m_1(B)=0.2, m_1(C)=0.3$;

Evidence 2: $m_2(A)=0, m_2(B)=0.9, m_2(C)=0.1$;

Evidence 3: $m_3(A)=0.55, m_3(B)=0.1, m_3(C)=0.35$;

Evidence 4: $m_4(A)=0.55, m_4(B)=0.1, m_4(C)=0.35$;

Evidence 5: $m_5(A)=0.55, m_5(B)=0.1, m_5(C)=0.35$;

It is obviously that most bodies of evidence support the identity A but the second body of evidence supports the identity B. In other words, they are in high degree of conflict. Fusion results offered by different combination rules are given in table 2.

Example 3 (two conflicting bodies of evidence): The basic belief assignments offered by five sensors are given as follows.

Evidence 1: $m_1(A)=0.5, m_1(B)=0.2, m_1(C)=0.3$;

Evidence 2: $m_2(A)=0, m_2(B)=0.9, m_2(C)=0.1$;

Evidence 3: $m_3(A)=0.3, m_3(B)=0.6, m_3(C)=0.1$;

Evidence 4: $m_4(A)=0.55, m_4(B)=0.1, m_4(C)=0.35$;

Evidence 5: $m_5(A)=0.55, m_5(B)=0.1, m_5(C)=0.35$;

As in example 2, most bodies of evidence support the identity A, but there are two bodies of evidence support the identity B. They are also in high degree of conflict. Fusion

results offered by different combination rules are given in table 3.

TABLE 1. FUSION RESULTS OF EXAMPLE 1

Fusion algorithm	Element of D^Θ	m_1m_2	$m_1m_2m_3$	$m_1m_2m_3m_4$	$m_1m_2m_3m_4m_5$
DST	A	0.75	0.8684	0.9213	0.9503
	B	0.1	0.0211	0.0041	0.0007
	C	0.15	0.1105	0.0746	0.049
PCR5	A	0.6529	0.7087	0.7373	0.7506
	B	0.1426	0.0602	0.0281	0.0201
	C	0.2045	0.2311	0.2346	0.2293
New self-adaptive combination rule	A	0.7289	0.8235	0.8598	0.8755
	B	0.1092	0.0303	0.0113	0.008
	C	0.1619	0.1462	0.1289	0.1165

TABLE 2. FUSION RESULTS OF EXAMPLE 2

Fusion algorithm	Element of D^Θ	m_1m_2	$m_1m_2m_3$	$m_1m_2m_3m_4$	$m_1m_2m_3m_4m_5$
DST	A	0	0	0	0
	B	0.8571	0.6316	0.3288	0.1228
	C	0.1429	0.3684	0.6712	0.8772
PCR5	A	0.2024	0.37	0.5101	0.6154
	B	0.6851	0.4482	0.258	0.1247
	C	0.1125	0.1818	0.2319	0.2599
New self-adaptive combination rule	A	0.1797	0.3727	0.5724	0.7366
	B	0.7044	0.4441	0.2068	0.0610
	C	0.1159	0.1832	0.2208	0.2024

TABLE 3. FUSION RESULTS OF EXAMPLE 3

Fusion algorithm	Element of D^Θ	m_1m_2	$m_1m_2m_3$	$m_1m_2m_3m_4$	$m_1m_2m_3m_4m_5$
DST	A	0	0	0	0
	B	0.8571	0.9730	0.9114	0.7461
	C	0.1429	0.0270	0.0886	0.2539
PCR5	A	0.2024	0.1920	0.3413	0.4827
	B	0.6851	0.7615	0.5116	0.3068
	C	0.1125	0.0465	0.1471	0.2105
New self-adaptive combination rule	A	0.1797	0.1246	0.2962	0.4904
	B	0.7044	0.8467	0.5791	0.3222
	C	0.1159	0.0287	0.1247	0.1874

Table 1 shows us that, DST is very suitable for fusing bodies of evidence in low degree of conflict. However, because PCR5 keeps the conflicting focal elements, the support degree of A is just 0.7506 when fusing the fifth body of evidence. It has been shown in Fig.3 that the degree of A offered by PCR5 is much less than the value offered by DST (0.9503). When applying the new self-adaptive algorithm, the support degree of A is 0.8755, which achieves a great improvement for that of PCR5. In other words, the new algorithm can fuse the sources of evidence in low degree of conflict well.

It can be seen from table 2 that, due to the high degree of conflict, the mass of A in fusion results by applying the DST rule is always 0. Obviously, it is illogical in real world. By using the PCR5 rule, one can make the right decision. However, we can see from Fig.4 that PCR5 rule could only offer slow convergence, while the new self-adaptive fusion algorithm not only overcomes the shortage of DST whose fusion result is illogical but also makes the right decision with quick convergence.

Table 3 gives the fusion results of example 3. Although there are two bodies of evidence which are in conflict with other evidence, our algorithm and PCR5 will give the right results. Besides, it can be seen from Fig.5 that, as the number of bodies of evidence increases, the new algorithm will get a better performance on convergence than PCR5. In conclusion, the new self-adaptive fusion algorithm is the best one among the three algorithms while dealing with high degree of conflict.

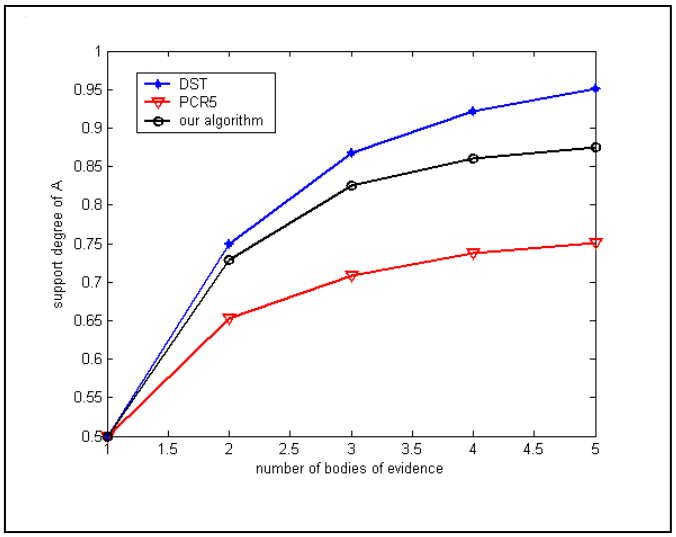


Fig.3. Support degree of A in Example 1

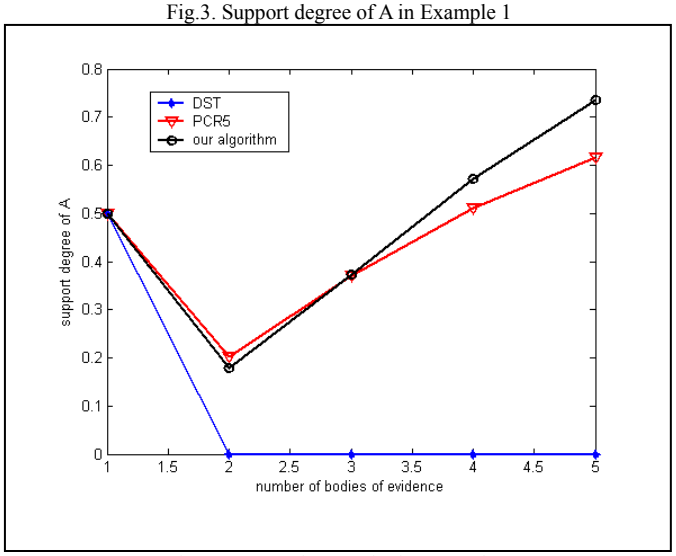


Fig.4. Support degree of A in Example 2

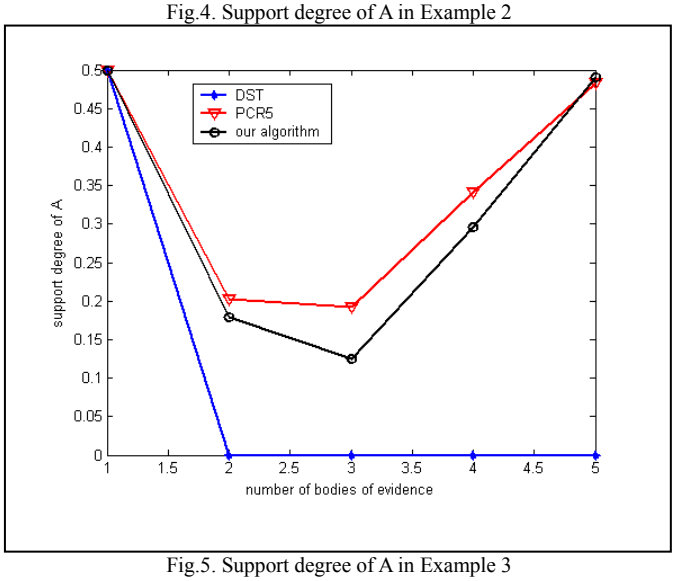


Fig.5. Support degree of A in Example 3

To sum up, the new self-adaptive algorithm can deal with high degree of conflict with a good performance on convergence.

V. CONCLUSION

Based on DST and DSmT, this paper proposes a new self-adaptive fusion algorithm. A controlling factor is introduced to avoid setting of the conflict threshold. Simulation results show that the new model can reach a preferable fusion result no matter the sources of evidence are in high degree of conflict or not. Furthermore, the new algorithm offers a quick convergence and it is more appropriate to be used in the real fusion system.

REFERENCES

[1] Laurence Cholvy. "Using logic to understand relations between DSmT and Dempster-Shafer theory," ECSQARU 2009, LNAI 5590, pp. 264–274, 2009.

[2] Yager. "On the Dempster-Shafer framework and new combination rules," Information System, vol. 41, no. 2, pp. 93-137, 1989.

[3] Murphy C K. "Combining belief functions when evidence conflicts," Decision Support Systems, vol.29, pp.1-9, 2000.

[4] Yong D, Kang S Y, Zhu Z F. "Combining belief functions based on distance of evidence," Decision Support Systems, vol.38, pp. 489-493, 2004.

[5] F. Smarandache, J. Dezert (Editors). "Advances and applications of DSmT for information fusion," Vol.1. American Research Press, Rehoboth, 2004.

[6] Weiru Liu. "Analyzing the degree of conflict among belief functions," Artificial Intelligence, vol. 170, pp. 909-924, 2006.

[7] Cuzzolin, F. A. "Geometric approach to the theory of evidence," IEEE, Transactions on System, Man, and Cybernetics-Part C: Applications and Reviews, vol. 38, no. 4, pp. 522-534, 2008.

[8] Ristic, B. and Smets, P. "The TBM global distance measure for the association of uncertain combat ID declaration," Information fusion, vol. 7, pp. 276-284, 2006.

[9] Tessem, B. "Approximations for efficient computation in the theory of evidence," Artificial Intelligence, vol.61, no. 2, pp. 315-329, 1993.

[10] A.L. Jousselme, D. Grenier, E. Bosse. "A new distance between two bodies of evidence," Information Fusion, vol. 2, pp. 91-101, 2001.

[11] F. Smarandache, J. Dezert (Editors). "Advances and applications of DSmT for information fusion," Vol.2, American Research Press, Rehoboth, 2006.

[12] Wen Jiang, Jingye Peng, Yong Deng. "New representation method of evidential conflict," Systems Engineering and Electronics, vol. 32, no.3, pp. 562-565, 2010 (in Chinese).

[13] Zhunga Liu, Yongmei Cheng, Quan Pan, Zhuang Miao. "Combination of weighted belief functions based on evidence distance and conflicting belief," Control Theory & Applications, vol.26, no. 12, pp.1439-1442, 2009 (in Chinese).

On the Quality of Optimal Assignment for Data Association

Jean Dezert
Kaouthar Benameur

Originally published as Dezert J., Benameur K., *On the quality of optimal assignment for data association*, in Proc. of Belief 2014 Conf. Oxford, UK, Sept. 26-29, 2014, and reprinted with permission.

Abstract. In this paper, we present a method based on belief functions to evaluate the quality of the optimal assignment solution of a classical association problem encountered in multiple target tracking applications. The purpose of this work is not to provide a new algorithm for solving the assignment problem, but a solution to estimate the quality of the individual associations (pairings) given in the optimal assignment solution. To the knowledge of authors, this problem has not been addressed so far in the literature and its solution may have practical aspects for improving the performances of multisensor-multitarget tracking systems.

Keywords: Data association; PCR6 rule; Belief function.

1 Introduction

Efficient algorithms for modern multisensor-multitarget tracking (MS-MTT) systems [1, 2] require to estimate and predict the states (position, velocity, etc) of the targets evolving in the surveillance area covered by the sensors. The estimations and the predictions are based on sensors measurements and dynamical models assumptions. In the monosensor context, MTT requires to solve the data association (DA) problem to associate the available measurements at a given time with the predicted states of the targets to update their tracks using filtering techniques (Kalman filter, Particle filter, etc). In the multisensor MTT context, we need to solve more difficult multi-dimensional assignment problems under constraints. Fortunately, efficient algorithms have been developed in operational research and tracking communities for formalizing and solving these optimal assignments problems. Several approaches based on different models can be used to establish rewards matrix, either based on the probabilistic framework [1, 3], or on the belief function (BF) framework [4–7]. In this paper, we do not focus on the construction of the rewards matrix[1], and our purpose is to provide a method to evaluate the quality (interpreted as a confidence score) of each association (pairing) provided in the optimal solution based on its consistency (stability) with respect to all the second best solutions.

[1] We assume that the rewards matrix is known and has been obtained by a method chosen by the user, either in the probabilistic or in the BF framework.

The simple DA problem under concern can be formulated as follows. We have $m > 1$ targets T_i $(i = 1, \ldots, m)$, and $n > 1$ measurements[2] z_j $(j = 1, \ldots, n)$ at a given time k, and a $m \times n$ rewards (gain/payoff) matrix $\boldsymbol{\Omega} = [\omega(i,j)]$ whose elements $\omega(i,j) \geq 0$ represent the payoff (usually homogeneous to the likelihood) of the association of target T_i with measurement z_j, denoted (T_i, z_j). The data association problem consists in finding the global optimal assignment of the targets to some measurements by maximizing[3] the overall gain in such a way that no more than one target is assigned to a measurement, and reciprocally.

Without loss of generality, we can assume $\omega(i,j) \geq 0$ because if some elements $\omega(i,j)$ of $\boldsymbol{\Omega}$ were negative, we can always add the same maximal negative value to all elements of $\boldsymbol{\Omega}$ to work with a new payoff matrix $\boldsymbol{\Omega}' = [\omega'(i,j)]$ having all elements $\omega'(i,j) \geq 0$, and we get the same optimal assignment solution with $\boldsymbol{\Omega}$ and with $\boldsymbol{\Omega}'$. Moreover, we can also assume, without loss of generality $m \leq n$, because otherwise we can always swap the roles of targets and measurements in the mathematical problem definition by working directly with $\boldsymbol{\Omega}^t$ instead, where the superscript t denotes the transposition of the matrix. The optimal assignment problem consists of finding the $m \times n$ binary association matrix $\mathbf{A} = [a(i,j)]$ which maximize the global rewards $R(\boldsymbol{\Omega}, \mathbf{A})$ given by

$$R(\boldsymbol{\Omega}, \mathbf{A}) \triangleq \sum_{i=1}^{m} \sum_{j=1}^{n} \omega(i,j) a(i,j) \tag{1}$$

$$\text{Subject to} \quad \begin{cases} \sum_{j=1}^{n} a(i,j) = 1 \quad (i = 1, \ldots, m) \\ \sum_{i=1}^{m} a(i,j) \leq 1 \quad (j = 1, \ldots, n) \\ a(i,j) \in \{0,1\} \quad (i = 1, \ldots, m \text{ and } j = 1, \ldots, n) \end{cases} \tag{2}$$

The association indicator value $a(i,j) = 1$ means that the corresponding target T_i and measurement z_j are associated, and $a(i,j) = 0$ means that they are not associated $(i = 1, \ldots, m$ and $j = 1, \ldots, n)$.

The solution of the optimal assignment problem stated in (1)–(2) is well reported in the literature and several efficient methods have been developed in the operational research and tracking communities to solve it. The most well-known algorithms are Kuhn-Munkres (a.k.as Hungarian) algorithm [8, 9] and its extension to rectangular matrices proposed by Bourgeois and Lassalle in [10], Jonker-Volgenant method [11], and Auction [12]. More sophisticated methods using Murty's method [13], and some variants [3, 14–19], are also able to provide not only the best assignment, but also the m-best assignments. We will not present in details all these classical methods because they have been already well reported in the literature [20, 21], and they are quite easily accessible on the

[2] In a multi-sensor context targets can be replaced by tracks provided by a given tracker associated with a type of sensor, and measurements can be replaced by another tracks set. In different contexts, possible equivalents are assigning personnel to jobs or assigning delivery trucks to locations.

[3] In some problems, $\boldsymbol{\Omega} = [\omega(i,j)]$ represents a cost matrix whose elements are the negative log-likelihood of association hypotheses. In this case, the data association problems consists in finding the best assignment that minimizes the overall cost.

web. In this paper, we want to provide a confidence level (i.e. a quality indicator) in the optimal data association solution. More precisely, we are searching an answer to the question: how to measure the quality of the pairings $a(i,j)=1$ provided in the optimal assignment solution $\mathbf{A}$? The necessity to establish a quality indicator is motivated by the following three main reasons:

1. In some practical tracking environment with the presence of clutter, some association decisions ($a(i,j)=1$) are doubtful. For these unreliable associations, it is better to wait for new information (measurements) instead of applying the hard data association decision, and making potentially serious association mistakes.
2. In some multisensor systems, it can be also important to save energy consumption for preserving a high autonomy capacities of the system. For this goal, only the most trustful specific associations provided in the optimal assignment have to be selected and used instead of all of them.
3. The best optimal assignment solution is not necessarily unique. In such situation, the establishment of quality indicators may help in selecting one particular optimal assignment solution among multiple possible choices.

Before presenting our solution in Section 2, one must recall that the best, as well as the 2nd-best, optimal assignment solutions are unfortunately not necessarily unique. Therefore, we must also take into account the possible multiplicity of assignments in the analysis of the problem. The multiplicity index of the best optimal assignment solution is denoted $\beta_1 \geq 1$, and the multiplicity index of the 2nd-best optimal assignment solution is denoted $\beta_2 \geq 1$, and we will denote the sets of corresponding assignment matrices by $\mathcal{A}_1 = \{\mathbf{A}_1^{(k_1)}, k_1 = 1\ldots,\beta_1\}$ and by $\mathcal{A}_2 = \{\mathbf{A}_2^{(k_2)}, k_2 = 1\ldots,\beta_2\}$. The next simple example illustrates a case with multiplicity of 2nd-best assignment solutions for the reward matrix $\boldsymbol{\Omega}_1$.

Example: $\beta_1 = 1$ and $\beta_2 = 4$ (i.e. no multiplicity of $\mathbf{A}_1$ and multiplicity of $\mathbf{A}_2$)

$$\boldsymbol{\Omega}_1 = \begin{bmatrix} 1 & 11 & 45 & 30 \\ 17 & 8 & 38 & 27 \\ 10 & 14 & 35 & 20 \end{bmatrix}$$

This reward matrix provides a unique best assignment $\mathbf{A}_1$ providing $R_1(\boldsymbol{\Omega}_1, \mathbf{A}_1) = 86$, and $\beta_2 = 4$ second-best assignment solutions providing $R_2(\boldsymbol{\Omega}_1, \mathbf{A}_2^{k_2}) = 82$ $(k_2 = 1,2,3,4)$ given by

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{A}_2^{k_2=1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ \mathbf{A}_2^{k_2=2} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ \mathbf{A}_2^{k_2=3} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix},\ \mathbf{A}_2^{k_2=4} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

2 Quality of the Associations of the Optimal Assignment

To establish the quality of the specific associations (pairings) (i,j) satisfying $a_1(i,j)=1$ belonging to the optimal assignment matrix $\mathbf{A}_1$, we propose to use both $\mathbf{A}_1$ and 2nd-best assignment solution $\mathbf{A}_2$. The basic idea is to compare the values $a_1(i,j)$ with $a_2(i,j)$ obtained in the best and in the 2nd-best assignments to identify the change (if any) of the optimal pairing (i,j). Our quality indicator will depend on both the stability of the pairing and its relative impact in the global reward. The proposed method works also when the 2nd-best assignment solution $\mathbf{A}_2$ is not unique (as in our example). The proposed method will also help to select the best (most trustful) optimal assignment in case of multiplicity of $\mathbf{A}_1$ matrices.

2.1 A Simplistic Method (Method I)

Before presenting our sophisticate method based on belief functions, let's first present a simplistic intuitive method (called Method I). For this, let's assume at first that $\mathbf{A}_1$ and $\mathbf{A}_2$ are unique (no multiplicity occurs). The simplistic method uses only the ratio of global rewards $\rho \triangleq R_2(\boldsymbol{\Omega},\mathbf{A}_2)/R_1(\boldsymbol{\Omega},\mathbf{A}_1)$ to measure the level of uncertainty in the change (if any) of pairing (i,j) provided in $\mathbf{A}_1$ and $\mathbf{A}_2$. More precisely, the quality (trustfulness) of pairings in an optimal assignment solution $\mathbf{A}_1$, denoted[4] $q_I(i,j)$, is simply defined as follows for $i=1,\ldots,m$ and $j=2,\ldots,n$:

$$q_I(i,j) \triangleq \begin{cases} 1, & \text{if } a_1(i,j)+a_2(i,j)=0 \\ 1-\rho & \text{if } a_1(i,j)+a_2(i,j)=1 \\ 1, & \text{if } a_1(i,j)+a_2(i,j)=2 \end{cases} \tag{3}$$

By adopting such definition, one commits the full confidence to the components (i,j) of $\mathbf{A}_1$ and $\mathbf{A}_2$ that perfectly match, and a lower confidence value (a lower quality) of $1-\rho$ to those that do not match. To take into account the eventual multiplicities (when $\beta_2>1$) of the 2nd-best assignment solutions $\mathbf{A}_2^{k_2}$, $k_2=1,2,\ldots,\beta_2$, we need to combine the $\mathbf{Q}_I(\mathbf{A}_1,\mathbf{A}_2^{k_2})$ values. Several methods can be used for this, in particular we can use either:

– **A weighted averaging approach**: The quality indicator component $q_I(i,j)$ is then obtained by averaging the qualities obtained from each comparison of $\mathbf{A}_1$ with $\mathbf{A}_2^{k_2}$. More precisely, one will take:

$$q_I(i,j) \triangleq \sum_{k_2=1}^{\beta_2} w(\mathbf{A}_2^{k_2}) q_I^{k_2}(i,j) \tag{4}$$

where $q_I^{k_2}(i,j)$ is defined as in (3) (with $a_2(i,j)$ replaced by $a_2^{k_2}(i,j)$ in the formula), and where $w(\mathbf{A}_2^{k_2})$ is a weighting factor in $[0,1]$, such that $\sum_{k_2=1}^{\beta_2} w(\mathbf{A}_2^{k_2})=1$. Since all assignments $\mathbf{A}_2^{k_2}$ have the same global reward value R_2, then we suggest to take $w(\mathbf{A}_2^{k_2})=1/\beta_2$. A more elaborate method

[4] The subscript I in $q_I(i,j)$ notation refers to Method I.

would consist to use the quality indicator of $\mathbf{A}_2^{k_2}$ based on the 3rd-best solution, which can be itself computed from the quality of the 3rd assignment solution based on the 4th-best solution, and so on by a similar mechanism. We however don't give more details on this due to space constraints.

– **A belief-based approach** (see [22] for basics on belief functions): A second method would express the quality by a belief interval $[q_I^{\min}(i,j), q_I^{\max}(i,j)]$ in $[0,1]$ instead of single real number $q_I(i,j)$ in $[0,1]$. More precisely, one can compute the belief and plausibility bounds of the quality by taking $q_I^{\min}(i,j) \equiv Bel(a_1(i,j)) = \min_{k_2} q_I^{k_2}(i,j)$ and $q_I^{\max}(i,j) \equiv Pl(a_1(i,j)) = \max_{k_2} q_I^{k_2}(i,j)$, with $q_I^{k_2}(i,j)$ given by (3) and $a_2(i,j)$ replaced by $a_2^{k_2}(i,j)$ in the formula. Hence for each association $a_1(i,j)$, one can define a basic belief assignment (BBA) $m_{ij}(.)$ on the frame of discernment $\Theta \triangleq \{T = \text{trustful}, \neg T = \text{not trustful}\}$, which will characterize the quality of the pairing (i,j) in the optimal assignment solution $\mathbf{A}_1$, as follows:

$$\begin{cases} m_{ij}(T) = q_I^{\min}(i,j) \\ m_{ij}(\neg T) = 1 - q_I^{\max}(i,j) \\ m_{ij}(T \cup \neg T) = q_I^{\max}(i,j) - q_I^{\min}(i,j) \end{cases} \tag{5}$$

Remark: In practice, only the pairings[5] (i,j) such that $a_1(i,j) = 1$ are useful in tracking algorithms to update the tracks. Therefore, we don't need to pay attention (compute and store) the qualities of components (i,j) such that $a_1(i,j) = 0$.

2.2 A More Sophisticate and Efficient Method (Method II)

The previous method can be easily applied in practice but it does not work very well because the quality indicator depends only on the ρ factor, which means that all mismatches between the best assignment $\mathbf{A}_1$ and the 2nd-best assignment solution $\mathbf{A}_2$ have their quality impacted in the same manner (they are all taken as $1-\rho$). As a simple example, if we consider the rewards matrix $\boldsymbol{\Omega}_1$ given in our example, we will have $\rho = R_2(\boldsymbol{\Omega}_1, \mathbf{A}_2^{k_2})/R_1(\boldsymbol{\Omega}_1, \mathbf{A}_1) = 82/86 \approx 0.95$, and we will get using method I with the weighting averaging approach (using same $w(\mathbf{A}_2^{k_2}) = 1/\beta_2 = 0.25$ for $k_2 = 1,2,3,4$) the following quality indicator matrix:

$$\mathbf{Q}_I(\mathbf{A}_1, \mathcal{A}_2) = \frac{1}{\beta_2} \sum_{k_2=1}^{\beta_2} \mathbf{Q}_I(\mathbf{A}_1, \mathbf{A}_2^{k_2}) = \begin{bmatrix} 1.0000 & 1.0000 & \mathbf{0.5233} & 0.5233 \\ 0.5233 & 1.0000 & 0.7616 & \mathbf{0.2849} \\ 0.7616 & \mathbf{0.2849} & 0.7616 & 0.7616 \end{bmatrix} \tag{6}$$

We observe that optimal pairings (2,4) and (3,2) get the same quality value 0.2849 with the method I (based on averaging), even if these pairings have different impacts in the global reward value, which is abnormal. If we use the method I with the belief interval measure based on (5), the situation is worst because the three optimal pairings (1,3), (2,4) and (3,2) will get exactly same belief interval values [0.0465,1]. To take into account, and in a better way, the

[5] given in the optimal solution found for example with Murty's algorithm.

reward values of each specific association given in the best assignment $\mathbf{A}_1$ and in the 2nd-best assignment $\mathbf{A}_2^{k_2}$, we propose to use the following construction of quality indicators depending on the type of matching (called Method II):

- When $a_1(i,j) = a_2^{k_2}(i,j) = 0$, one has full agreement on "non-association" (T_i, z_j) in $\mathbf{A}_1$ and in $\mathbf{A}_2^{k_2}$ and this non-association (T_i, z_j) has no impact on the global rewards values $R_1(\boldsymbol{\Omega}, \mathbf{A}_1)$ and $R_2(\boldsymbol{\Omega}, \mathbf{A}_2^{k_2})$, and it will be useless. Therefore, we can set its quality arbitrarily to $q_{II}^{k_2}(i,j) = 1$.
- When $a_1(i,j) = a_2^{k_2}(i,j) = 1$, one has a full agreement on the association (T_i, z_j) in $\mathbf{A}_1$ and in $\mathbf{A}_2^{k_2}$ and this association (T_i, z_j) has different impacts in the global rewards values $R_1(\boldsymbol{\Omega}, \mathbf{A}_1)$ and $R_2(\boldsymbol{\Omega}, \mathbf{A}_2^{k_2})$. To qualify the quality of this matching association (T_i, z_j), we define the two BBA's on $X \triangleq (T_i, z_j)$ and $X \cup \neg X$ (the ignorance), for $s = 1, 2$:

$$\begin{cases} m_s(X) = a_s(i,j) \cdot \omega(i,j) / R_s(\boldsymbol{\Omega}, \mathbf{A}_s) \\ m_s(X \cup \neg X) = 1 - m_s(X) \end{cases} \tag{7}$$

  Applying the conjunctive rule of fusion, we get

$$\begin{cases} m(X) = m_1(X)m_2(X) + m_1(X)m_2(X \cup \neg X) + m_1(X \cup \neg X)m_2(X) \\ m(X \cup \neg X) = m_1(X \cup \neg X)m_2(X \cup \neg X) \end{cases} \tag{8}$$

  Applying the pignistic transformation[6] [24], we get finally $BetP(X) = m(X) + \frac{1}{2} \cdot m(X \cup \neg X)$ and $BetP(\neg X) = \frac{1}{2} \cdot m(X \cup \neg X)$. Therefore, we choose the quality indicator as $q_{II}^{k_2}(i,j) = BetP(X)$.
- When $a_1(i,j) = 1$ and $a_2^{k_2}(i,j) = 0$, one has a disagreement (conflict) on the association (T_i, z_j) in $\mathbf{A}_1$ and in (T_i, z_{j_2}) in $\mathbf{A}_2^{k_2}$, where j_2 is the measurement index such that $a_2(i, j_2) = 1$. To qualify the quality of this non-matching association (T_i, z_j), we define the two following basic belief assignments (BBA's) of the propositions $X \triangleq (T_i, z_j)$ and $Y \triangleq (T_i, z_{j_2})$

$$\begin{cases} m_1(X) = a_1(i,j) \cdot \frac{\omega(i,j)}{R_1(\boldsymbol{\Omega}, \mathbf{A}_1)} \\ m_1(X \cup Y) = 1 - m_1(X) \end{cases} \quad \text{and} \quad \begin{cases} m_2(Y) = a_2(i,j_2) \cdot \frac{\omega(i,j_2)}{R_2(\boldsymbol{\Omega}, \mathbf{A}_2^{k_2})} \\ m_2(X \cup Y) = 1 - m_2(Y) \end{cases} \tag{9}$$

  Applying the conjunctive rule, we get $m(X \cap Y = \emptyset) = m_1(X)m_2(Y)$ and

$$\begin{cases} m(X) = m_1(X)m_2(X \cup Y) \\ m(Y) = m_1(X \cup Y)m_2(Y) \\ m(X \cup Y) = m_1(X \cup Y)m_2(X \cup Y) \end{cases} \tag{10}$$

  Because we need to work with a normalized combined BBA, we can choose different rules of combination (Dempster-Shafer's, Dubois-Prade's,Yager's

[6] We have chosen here BetP for its simplicity and because it is widely known, but DSmP could be used instead for expecting better performances [23].

rule [23], etc). In this work, we recommend the Proportional Conflict Redistribution rule no. 6 (PCR6), proposed originally in DSmT framework [23], because it has been proved very efficient in practice. So, we get with PCR6:

$$\begin{cases} m(X) = m_1(X)m_2(X \cup Y) + m_1(X) \cdot \frac{m_1(X)m_2(Y)}{m_1(X)+m_2(Y)} \\ m(Y) = m_1(X \cup Y)m_2(Y) + m_2(X) \cdot \frac{m_1(X)m_2(Y)}{m_1(X)+m_2(Y)} \\ m(X \cup Y) = m_1(X \cup Y)m_2(X \cup Y) \end{cases} \quad (11)$$

Applying the pignistic transformation, we get finally $BetP(X) = m(X) + \frac{1}{2} \cdot m(X \cup Y)$ and $BetP(Y) = m(Y) + \frac{1}{2} \cdot m(X \cup Y)$. Therefore, we choose the quality indicators as follows: $q_{II}^{k_2}(i,j) = BetP(X)$, and $q_{II}^{k_2}(i,j_2) = BetP(Y)$.

The absolute quality factor $Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2})$ of the optimal assignment given in $\mathbf{A}_1$ conditioned by $\mathbf{A}_2^{k_2}$, for any $k_2 \in \{1, 2, \ldots, \beta_2\}$ is defined as

$$Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2}) \triangleq \sum_{i=1}^{m} \sum_{j=1}^{n} a_1(i,j) q_{II}^{k_2}(i,j) \quad (12)$$

Example (continued): If we apply the Method II (using PCR6 fusion rule) to the rewards matrix $\boldsymbol{\Omega}_1$, then we will get the following quality matrix (using weighted averaging approach)

$$\mathbf{Q}_{II}(\mathbf{A}_1, \mathcal{A}_2) = \frac{1}{\beta_2} \sum_{k_2=1}^{\beta_2} \mathbf{Q}_{II}(\mathbf{A}_1, \mathbf{A}_2^{k_2}) = \begin{bmatrix} 1.0000 & 1.0000 & \mathbf{0.7440} & 0.7022 \\ 0.7200 & 1.0000 & 0.8972 & \mathbf{0.5753} \\ 0.8695 & \mathbf{0.4957} & 0.9119 & 0.8861 \end{bmatrix}$$

with the absolute quality factors $Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=1}) \approx 1.66$, $Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=2}) \approx 1.91$, $Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=3}) \approx 2.19$, $Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=4}) \approx 1.51$. Naturally, we get

$$Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=3}) > Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=2}) > Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=1}) > Q_{abs}(\mathbf{A}_1, \mathbf{A}_2^{k_2=4})$$

because $\mathbf{A}_1$ has more matching pairings with $\mathbf{A}_2^{k_2=3}$ than with other 2nd-best assignment $\mathbf{A}_2^{k_2}$ ($k_2 \neq 3$), and those pairings have also the strongest impacts in the global reward value. One sees that the quality matrix $\mathbf{Q}_{II}$ differentiates the qualities of each pairing in the optimal assignment $\mathbf{A}_1$ as expected (contrariwise to Method I). Clearly, with Method I we obtain the same quality indicator value 0.2849 for the specific associations (2,4) and (3,2) which seems intuitively not very reasonable because the specific rewards of these associations impact differently the global rewards result. If the method II based on the belief interval measure computed from (5) is preferred[7], we will get respectively for the three optimal pairings (1,3), (2,4) and (3,2) the three distinct belief interval [0.5956,0.8924], [0.4113,0.7699] and [0.3524,0.6529]. These belief intervals show that the ordering of quality of optimal pairings (based either on the lower bound, or on the upper bound of belief interval) is consistent with the ordering of quality of optimal pairings in $\mathbf{Q}_{II}(\mathbf{A}_1, \mathcal{A}_2)$ computed with the averaging approach. Method II provides a better effective and comprehensive solution to estimate the quality of each specific association provided in the optimal assignment solution $\mathbf{A}_1$.

[7] just in case of multiplicity of second best assignments.

3 Conclusion

In this paper we have proposed a method based on belief functions for establishing the quality of pairings belonging to the optimal data association (or assignment) solution provided by a chosen algorithm. Our method is independent of the choice of the algorithm used in finding the optimal assignment solution, and, in case of multiple optimal solutions, it provides also a way to select the best optimal assignment solution (the one having the highest absolute quality factor). The method developed in this paper is general in the sense that it can be applied to different types of association problems corresponding to different sets of constraints. This method can be extended to SD-assignment problems. The application of this approach in a realistic multi-target tracking context is under investigations and will be reported in a forthcoming publication if possible.

References

1. Bar-Shalom, Y., Willet, P.K., Tian, X.: Tracking and Data Fusion: A Handbook of Algorithms. YBS Publishing, Storrs, CT, USA (2011)
2. Hall, D.L., Chong C.Y., Llinas, Liggins II, M.: Distributed Data Fusion for Network-Centric Operations. CRC Press (2013)
3. He, X., Tharmarasa, R., Pelletier, M., Kirubarajan, T.: Accurate Murty's Algorithm for Multitarget Top Hypothesis. In: Proc. of Fusion 2011, Chicago, USA (2011)
4. Dezert, J., Smarandache, F., Tchamova, A.: On the Blackman's Association Problem. In: Proc. of Fusion 2003, Cairns, Australia (2003)
5. Tchamova, A., Dezert, J., Semerdjiev, Tz., Konstantinova, P.: Target Tracking with Generalized Data Association Based on the General DSm Rule of Combination. In: Proc. of Fusion 2004, Stockholm, Sweden (2004)
6. Dezert, J., Tchamova, A., Semerdjiev, T., Konstantinova, P.: Performance Evaluation of Fusion Rules For Multitarget Tracking in Clutter Based on Generalized Data Association. In: Proc. of Fusion 2005, Philadelphia, USA (2005)
7. Denœux T., El Zoghby N., Cherfaoui V., Jouglet A.: Optimal Object Association in the Dempster- Shafer Framework. To appear in IEEE Trans. on Cybern. (2014)
8. Kuhn, H.W.: The Hungarian Method for the Assignment Problem. Naval Research Logistic Quarterly, vol. 2, 83–97 (1955)
9. Munkres, J.: Algorithms for the Assignment and Transportation Problems. Journal of the Society of Industrial and Applied Mathematics, vol. 5 (1), 32–38 (1957)
10. Bourgeois, F., Lassalle, J.C.: An Extension of the Munkres Algorithm for the Assignment Problem to Rectangular Matrices. Comm. of the ACM, vol. 14 (12), 802–804, Dec. (1971)
11. Jonker, R., Volgenant, A.: A shortest augmenting path algorithm for dense and sparse linear assignment problems. J. of Comp., vol. 38 (4), 325–340 (1987)
12. Bertsekas, D.: The auction algorithm: A Distributed Relaxation Method for the Assignment Problem. Annals of Operations Research, vol. 14 (1), 105–123 (1988)
13. Murty, K.G.: An Algorithm for Ranking all the Assignments in Order of Increasing Cost. Operations Research, vol. 16 (3), 682–687 (1968)
14. Chegireddy, C.R., Hamacher, H.W.: Algorithms for Finding K-best Perfect Matching. Discrete Applied Mathematics, vol. 18, 155–165 (1987)

15. Danchick, R., Newnam, G.E.: A Fast Method for Finding the Exact N-best Hypotheses for Multitarget Tracking. IEEE Trans. on AES, vol. 29 (2), 555–560 (1993)
16. Miller, M.L., Stone, H.S., Cox, I.J.: Optimizing Murty's Ranked Assignment Method. IEEE Trans. on AES, vol. 33 (3), 851–862, July (1997)
17. Pascoal, M., Captivo, M.E., Climaco, J.: A Note on a New Variant of Murtys Ranking Assignments Algorithm. 4OR Journal, vol. 1, 243–255 (2003)
18. Ding, Z., Vandervies, D.: A Modified Murty's Algorithm for Multiple Hypothesis Tracking. In: SPIE Sign. and Data Proc. of Small Targets, vol. 6236, (2006)
19. Fortunato, E., et al.: Generalized Murty's Algorithm with Application to Multiple Hypothesis Tracking. In: Proc. of Fusion 2007, 1–8, Québec, July 9–12 (2007)
20. Hamacher, H.W., Queyranne, M.: K-best Solutions to Combinatorial Optimization Problems. Annals of Operations Research, no. 4 , 123–143 (1985/6)
21. Dell'Amico, M., Toth, P.: Algorithms and Codes for Dense Assignment Problems: the State of the Art. Discrete Applied Mathematics, vol. 100, 17–48 (2000)
22. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, Princeton, New Jersey, USA (1976)
23. Smarandache, F., Dezert, J.: Advances and Applications of DSmT for Information Fusion, Volumes 1, 2 & 3. ARP (2004–2009)
http://fs.gallup.unm.edu/DSmT.htm
24. Smets, P., Kennes, R.: The Transferable Belief Model. Artificial Intelligence, vol. 66 (2), 191–234 (1994)

Comparison of Identity Fusion Algorithms Using Estimations of Confusion Matrices

G. Golino
A. Graziano
A. Farina
W. Mellano
F. Ciaramaglia

Originally presented at the 17th International Conference on Information Fusion, Salamanca, 7-10 July 2014, as: G. Golino, A. Graziano, A. Farina, W. Mellano, and F. Ciaramaglia - *Comparison of identity fusion algorithms using estimations of confusion matrices.* Printed with authors' permission.

***Abstract*— Scope of this paper is to investigate the performances of different identity declaration fusion algorithms in terms of probability of correct classification, supposing that the information for combination of the inferences from the different classifier is affected by measurement errors. In particular, these information have been assumed to be provided in the form of confusion matrices. Six identity fusion algorithms from literature with different complexity have been included in the comparison: heuristic methods such as voting and Borda Count, Bayes' and Dempster-Shafer's methods and the Proportional Redistribution Rule n° 1 in the Dempster-Shafer's framework.**

***Keywords*—target classification, identity fusion, confusion matrix.**

I. INTRODUCTION

In a multi-sensor system the target classification performance can be improved by suitably combining the inferences generated by the autonomous classifiers of the single sensors (identity declaration fusion [1]). For this purpose it is desirable to use the available information about the classification performances of the single sensors. The confusion matrix, whose elements correspond to the likelihood of the different involved classes, is a compact and detailed way of representing the classification performance, from which the Probability of correct classification (Pcc) and the probability relative to the various misclassification errors can be derived. In particular the elements of the confusion matrix can be used to maximize the a-posterior Pcc according to Bayes' theory. In this case, if the numerical values of the confusion matrix were errorless, the performance of the identity fusion would be optimal. However, in practice, these values are estimated and affected by errors. In these conditions, the Bayes' rule does not always produce best results. In particular, in presence of strong-conflicting inferences and estimation errors, the application of Bayes' rule can be not effective. It can be better to apply simpler combination rules as some heuristic methods that are more robust to errors.
Dempster-Shafer's theory has been presented as a generalization of Bayes' theory in [2]. A recent work has disputed this claim, limiting its correctness to the case of uniform a-prior probabilities [3]. The problem of the Dempster-Shafer rule (and of Bayes' rule) in presence of conflicting inferences has been pointed out by the well-known Zadeh's paradox [4]. In this paper the performances of different algorithms that use estimated confusion matrices affected by errors to combine the inferences from the single classifiers are investigated. In particular the heuristic methods based on voting and on ranking (Borda count) are compared with the methods using Bayes' and Dempster-Shafer's rules. Moreover the effectiveness of the redistribution of the conflicting masses preserving the Dempster-Shafer framework, like Proportional Redistribution Rules (PCR) is evaluated.

The paper is organized as follows:

- in section II the identity fusion algorithms considered in this paper are briefly described;
- in section III, four simple but representative identity fusion problems are introduced as study cases and the corresponding results using different mean values of the estimation errors of the confusion matrices are reported and commented;
- section IV gives the conclusions.

II. IDENTITY FUSION ALGORITHMS

The algorithms for the identity fusion considered in this paper are:

- Majority Voting (MV),
- Weighted Voting (WV),
- Borda count,
- Bayes' rule,
- Dempster-Shafer's (D-S) rule with the following basic belief mass assignment: "q-least commitment",
- Proportional Redistribution Rule n°1 (PCR1) with the following basic belief mass assignment: "q-least commitment".

A brief description of the fusion algorithms follows.

Majority voting [5] is the simplest method for the combination of inferences: each inferred class corresponds to a single vote and the selected class after fusion is the most voted class: all the inferences matter the same. In the modified weighted version, the different votes are weighted by the estimated Pcc of the voter/classifier.

Voting methods use only the top choice of each classifier, but secondary choices often contain near misses that should not be overlooked. The Borda count [5] is a method in which

classes are ranked in order of preference; it gives each class a certain number of points corresponding to the position in which it is ranked by each classifier. The class with the highest scoring is then selected after fusion.

Bayes' theorem [1,4] links the degree of belief in a proposition before and after accounting for evidence (a-priori and a-posteriori probabilities). The a-posteriori probability of a combination of two or more evidences is obtained by the multiplication of the likelihoods of the single evidences (the independence of the evidences is assumed).

Dempster–Shafer theory [1,4 and 5] allows to specify a degree of ignorance instead of being forced to supply probabilities that add to unity. In this formalism a degree of belief (also referred to as a Basic Belief Mass - BBM) is used rather than a Bayesian probability distribution. BBM values are assigned to sets of possibilities (union of one or more classes) rather than to a single class, probability is instead represented by intervals that are lower-bounded by the value "belief" (or "support") and upper-bounded by the value "plausibility". BBM values from different sources can be combined with Dempster-Shafer's rule of combination, assuming independent belief sources. There are more than one possible assignment for transforming probabilities into BBMs [7,8 and 9]. The "q-least commitment" basic belief mass assignment (that corresponds to the maximum compatible degree of ignorance) has been considered in this paper to transform the CM (Confusion Matrix) likelihoods and the a-priori probabilities into BBM values.

Proportional Redistribution Rules (PCR) is a family of fusion rules for the combination of uncertain information allowing to deal with highly conflicting sources. The PCR rules can be used as alternatives to the Dempster-Shafer's combination rule. Six PCR rules (PCR1-PCR6) have been defined [10,11 and 12]: from PCR1 up to PCR6 one increases in one hand the complexity of the rules, but in other hand one improves the accuracy of the redistribution of conflicting masses. The basic common principle of PCR rules is to redistribute the conflicting mass proportionally with some functions depending on the sum of the masses assigned by the single inferences. PCR1 is the least accurate combination rule of the PCR family, but it is the simplest to implement and it has been considered in this paper. PCR2-6 implementations are significantly more complex because the conflicting mass is redistributed only to the non-empty set that are involved in the conflict (extra computer memory is needed to keep track of the conflicting hypotheses and extra computation load is needed for combining them). A particular interesting action point for further investigation would be testing the most efficient PCR rule (PCR6) [12].

A. Combination rules

In this section, the rules of Bayes, Dempster-Shafer and PCR1 for the combination of two classifiers are briefly recalled. For further details and the generalization of the rules with more than two classifiers, see references [1], [4], [10] and [11]. Voting and Borda count combinations are not considered here because they consists simply in the sums of respectively the votes and the ranks.

Let consider a set Ω of possible exhaustive and mutually exclusive classes C_k, with N being the cardinality of this set:

$$\Omega = \{C_1, C_2, \cdots, C_N\} \tag{1}$$

Let suppose that the independent classifiers 1 and 2 infer respectively the classes C_i and C_j; the a-posterior probability $P_{12}(C_k / A \cap B)$ of inferring the class C_k resulting from Bayes' rule of combination is:

$$P_{12}(C_k / C_i \cap C_j) = \frac{P_1(C_i/C_k) \cdot P_2(C_j/C_k) \cdot P_0(C_k)}{\sum_{h=1}^{N} P_1(C_i/C_h) \cdot P_2(C_j/C_h) \cdot P_0(C_h)} \tag{2}$$

where:

- $P_0(\cdot)$ is the a-prior probabilities (without any information obtained by previous classifications) of the considered class;
- $P_1(C_y/C_x)$, $P_2(C_y/C_x)$ are the probabilities that classifiers 1 and 2 infer the class C_y assuming that the true class is C_x (likelihoods).

Let consider the power set 2^{Ω} of Ω as the set whose elements are all the possible subsets of Ω:

$$2^{\Omega} = \{F_k : F_k \subseteq \Omega\} = \{\emptyset, C_1, C_2, \cdots, C_N, C_1 \cap C_2, \cdots, C_{N-1} \cap C_N, \cdots, \Omega\} \tag{3}$$

where $\emptyset$ is the empty set. The cardinality of 2^{Ω} is 2^N.

Let suppose that the independent classifiers 1 and 2 assign respectively BBMs $m_1(\cdot)$ and $m_2(\cdot)$ to the elements included in the power set 2^{Ω}; the combination BBM $m_{12}(F_k)$ of F_k resulting from Dempster-Shafer's rule of combination is:

$$m_{12}(F_k) = \sum_{i,j / F_k = F_i \cap F_j} \frac{m_1(F_i) \cdot m_2(F_j)}{1 - m_c} \tag{4}$$

where m_c is the global conflicting mass, defined as follow

$$m_c = \sum_{p,q / F_p \cap F_q = \emptyset} m_1(F_p) \cdot m_2(F_q) \tag{5}$$

In the case of the PCR1 rule the combination BBM $m_{12}(F_k)$ of F_k is instead:

$$m_{12}(F_k) = \sum_{i,j / F_i \cap F_j = F_k \neq \emptyset} m_1(F_i) \cdot m_2(F_j) + \frac{m_1(F_k) + m_2(F_k)}{\sum_{h=1}^{2^N} (m_1(F_h) + m_2(F_h))} \cdot m_c \tag{6}$$

III. STUDY CASES

Four simple but representative study cases (three different classifiers for a three classes problem) have been investigated, as follows:

- complementary confusion matrices,
- supplementary confusion matrices,
- complementary conflicting confusion matrices,
- supplementary conflicting confusion matrices.

By complementary CMs it has been meant that the single classifiers show a complementary expertise in the recognition of the different classes. By supplementary CMs the different classifiers show similar behaviors. By conflicting CMs a possible overestimation of the performance of the single classifiers can make harder an effective combination of the contradictory inferences from different classifiers when they occur. A quantitative definition of complementary and supplementary CM can be found in [13].

In the following sub-sections, the estimated confusion matrices that have been selected for the four study cases are reported. The columns of the matrices represent the true classes, while the rows correspond to the inferred classes, so the element (k,h) of a matrix is an estimation of the probability of declaring k^{th} class when the true class is the h^{th} one:

$$\hat{M}_i(k,h) = \hat{P}_i(D \equiv k / T \equiv h) \qquad (7)$$

The a-prior probabilities of the different classes are assumed equal. A block diagram of the fusion system is shown in fig. 1.

The performances of the six algorithms in correspondence of the identity fusion of six inferences (two independent inferences for each classifier) have been considered. The performances of the different algorithms have been computed with 1000 Monte Carlo trials, each generating independent samples of the true confusion matrices and a-prior probabilities.

The results of the Monte Carlo trials are represented by the curves corresponding to Empirical Cumulative Distribution Function (ECDF) versus the Pcc. The x-axis values of Pcc have been computed exactly, that is the contribution of all the possible permutations of the single sensor inferences has been considered.

A. Complementary confusion matrices

The following confusion matrices corresponding to three different classifiers have been considered:

$$\hat{M}_1 = \begin{bmatrix} 0.90 & 0.30 & 0.30 \\ 0.10 & 0.40 & 0.30 \\ 0 & 0.30 & 0.40 \end{bmatrix}$$
$$\hat{M}_2 = \begin{bmatrix} 0.40 & 0.30 & 0.10 \\ 0.30 & 0.40 & 0 \\ 0.30 & 0.30 & 0.90 \end{bmatrix} \qquad (8)$$
$$\hat{M}_3 = \begin{bmatrix} 0.40 & 0 & 0.30 \\ 0.30 & 0.90 & 0.30 \\ 0.30 & 0.10 & 0.40 \end{bmatrix}$$

The performance is dependent on the true target class (class 1, 2 or 3):

- the first classifier identifies correctly targets belonging to class 1 (on average it makes only one mistake in ten of its inferences), while it almost randomly infers in correspondence of targets belonging to class 2 or class 3,
- the second classifier identifies correctly targets belonging to class 3 (on average it makes only one mistake in ten of its inferences), while it almost randomly infers in correspondence of targets belonging to class 1 or class 2,
- the third classifier identifies correctly targets belonging to class 2 (on average it makes only one mistake in ten of its inferences), while it almost randomly infers in correspondence of targets belonging to class 1 or class 2.

The performances of the six different algorithms are reported in the fig. 2 and 3 in correspondence of an estimation of the confusion matrix by using respectively 30 and 10 independent samples for each class. The performances of the single classifiers correspond to the dotted curves (indicated as C1, C2 and C3 in the legends of the figures). Bayes' rule, Dempster-Shafer's rule, Borda count and PCR1 give similar results, the performance of PCR1 is barely the best. The voting algorithms present significantly worse performance.

B. Supplementary confusion matrices

The following confusion matrices corresponding to three different classifiers have been considered:

$$\hat{M}_1 = \begin{bmatrix} 0.70 & 0.10 & 0.20 \\ 0.20 & 0.70 & 0.10 \\ 0.10 & 0.20 & 0.70 \end{bmatrix}$$
$$\hat{M}_2 = \begin{bmatrix} 0.80 & 0.10 & 0.10 \\ 0.10 & 0.80 & 0.10 \\ 0.10 & 0.10 & 0.80 \end{bmatrix} \quad (9)$$
$$\hat{M}_3 = \begin{bmatrix} 0.60 & 0.20 & 0.20 \\ 0.20 & 0.60 & 0.20 \\ 0.20 & 0.20 & 0.60 \end{bmatrix}$$

In the second example, three classifiers with supplementary confusion matrices have been selected. A single classifier can recognize all the three classes with the same accuracy, but the accuracy differs from classifier to classifier:

- the first classifier has an estimated probability of correct classification equal to 70%,
- the second classifier has an estimated probability of correct classification equal to 80%,
- the third classifier has an estimated probability of correct classification equal to 60%.

The performances of the six different algorithms are reported in the fig. 4 and 5 in correspondence of an estimation of the confusion matrix by using respectively 30 and 10 independent samples for each class. The performances of the single classifiers correspond to the dotted curves (indicated as C1, C2 and C3 in the legends of the figures). All the algorithms give comparable performance. PCR1 and Bayes' rule performance are exactly the same and they are slightly better than the others, weighted voting performs better than Borda count and majority voting.

C. Complementary conflicting confusion matrices

The following confusion matrices corresponding to three different classifiers have been considered:

$$\hat{M}_1 = \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 0.60 & 0.40 \\ 0.00 & 0.40 & 0.60 \end{bmatrix}$$
$$\hat{M}_2 = \begin{bmatrix} 0.60 & 0.00 & 0.40 \\ 0.00 & 1.00 & 0.00 \\ 0.40 & 0.00 & 0.60 \end{bmatrix} \quad (10)$$
$$\hat{M}_3 = \begin{bmatrix} 0.60 & 0.40 & 0.00 \\ 0.40 & 0.60 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{bmatrix}$$

In this third example, the three classifiers can be affected by conflicting inferences. As consequence, the application of Bayes' rule to fusion leads to severe performance degradation with respect to heuristic methods. The problem arises from an overestimation of the performance of the single classifiers.

The performances of the six different algorithms are reported in the fig. 6 and 7 in correspondence of an estimation of the confusion matrix by using respectively 30 and 10 independent samples for each class. The performances of the single classifiers correspond to the dotted curves (indicated as C1, C2 and C3 in the legends of the figures). PCR1 gives the best result that is slight better than Borda count. Majority and weighted voting have coincident performance that are significantly worse than the one of PCR1. The performance of Bayes' rule and Dempster-Shafer rule are perfectly coincident and worse than all the others because of the presence of conflicting inferences.

D. Supplementary conflicting confusion matrices

The following confusion matrices corresponding to three different classifiers have been considered:

$$\hat{M}_1 = \hat{M}_2 = \hat{M}_3 = \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 0.00 \\ 0.00 & 0.00 & 1.00 \end{bmatrix} \quad (11)$$

In the forth example, three classifiers with identity confusion matrices as estimations have been selected: according to these estimations the single classifier is never wrong. If the classifiers disagree on the inferred class, the Bayes' rule of fusion leads to severe performance degradation with respect to heuristic methods.

The performances of the six different algorithms are reported in the fig. 8 and 9 in correspondence of an estimation of the confusion matrix by using respectively 30 and 10 independent samples for each class. The performance of the single classifiers correspond to the dotted curves (indicated as C1, C2 and C3 in the legends of the figures). The performances of the voting algorithms, Borda count and PCR1 are perfectly coincident and near to 100%. The performances of Bayes' rule and Dempster-Shafer rule are perfectly coincident and much worse than all the others because of the presence of conflicting inferences, even much worse than the performance of the single classifiers.

E. Summary results

In table I the average (over the 1000 Monte Carlo trials) Pcc is reported for all the investigated study cases. It has been reported also an intermediate case where 60 samples (20 samples per class) for the estimation of each confusion matrix have been considered. It can be noted than PCR1 always brings the highest Pcc of all the six considered combination rules.

IV. CONCLUSIONS

Simulation results show that in the considered study cases, the algorithm using the PCR1 rule of combination brings the best performance of all the six considered alternatives and definitely overcomes Bayes' and D-S's rules in the cases where the probability of conflicts between the inferences is high. This performance difference increases with the decrease of the number of samples used for the estimation of the confusion matrices. This behavior is a consequence of the poor performance of the latter two combination methods in presence of conflicting inferences from the different classifier, as claimed by the Zadeh's paradox. In two investigated study cases with conflicting confusion matrices Bayes' and Dempster-Shafer rules perform even worse than heuristic approaches.

In the cases where the conflict is less likely probable the performance of the PCR1 is comparable with the ones of Bayes' and D-S's rules (the same or slightly better).

The implementation of PCR1 slightly increases the computational complexity of D-S's rule. Future work may be addressed to the comparison of the performance resulting by the application of more complex PCR rules to the inferences of classifiers whose accuracies are represented by conflicting confusion matrices.

REFERENCES

[1] D. L. Hall, S. A. H. McMullen, "Mathematical Techniques in Multisensor Data Fusion", 2nd edition, Artech House, 2004.

[2] A. Dempster, "A generalization of bayesian inference", Journal of the Royal Statistical Society, series B, vol. 30, n° 2, pp. 205-247, 1968.

[3] J. Dezert , A. Tchamova, D. Han, J.M: Tacnet, "Why Dempster's fusion rule is not a generalization of Bayes fusion rule", Proc. of Fusion 2013 - 16th Int. Conference on Information Fusion, Istanbul, Turkey, July 9-12, 2013.

[4] A. Benavoli, L. Chisci, B. Ristic, A. Graziano, A. Farina, "Reasoning under uncertainty: from Bayes to Valuation Based Systems. Application to target classification and threat evaluation", Selex Sistemi Integrati, Università degli Studi di Firenze, September 2007.

[5] J. Llinas, "Fusion-based methods for target identification in the absence of quantitative classifier confidence", SPIE Signal Processing, Sensor Fusion and Target Recognition VI Conference, Orlando, FL, April 1997.

[6] A. Benavoli, B. Ristic, A. Farina, M. Oxenham, and L. Chisci, "An Application of Evidential Networks to Threat Assessment". IEEE Trans. of Aerospace and Electronic Systems, vol. 45, n° 2, pp. 620-639, April 2009.

[7] P. Smets, "Quantified epistemic possibility theory seen as a hyper cautious transferable belief model", Rencontres Francophones sur la Logique Floue et ses Applications, pages 343-353, Cepadues Editions, October 2000.

[8] P. Smets, "The transferable belief model for quantified belief representation", Handbook of Defeasible Reasoning and Uncertainty Management Systems, vol. 1, pp. 267-301, Kluwer, Dordrecht, The Netherlands, 1998.

[9] B. Ristic, P. Smets, "Target classification approach based on belief function theory", IEEE Trans. of Aerospace and Electronic Systems, vol. 41, n° 2, pp. 574-583, April 2005.

[10] J. Dezert, F. Smarandache, "Information fusion based on new proportional conflict redistribution rules" , Proc. of Fusion 2005 - 8th Int. Conference on Information Fusion, Philadelphia, PA, USA, June 27 - July 1, 2005.

[11] J. Dezert, F. Smarandache (editors), "Advances and Applications of DSmT for Information Fusion (Collected works), second volume", American Research Press, 2006.

[12] J. Dezert, F. Smarandache, "On the consistency of PCR6 with the averaging rule and its application to probability estimation", Proc. of Fusion 2013 - 16th Int. Conference on Information Fusion, Istanbul, Turkey, July 9-12, 2013.

[13] S. R. M. Prasanna, B. Yegnanarayana, J. P. Pinto, H. Hermansky, "Analysis of Confusion Matrix to Combine Evidence for Phoneme Recognition", Idiap Resarch Institute, Switzerland, 2007.

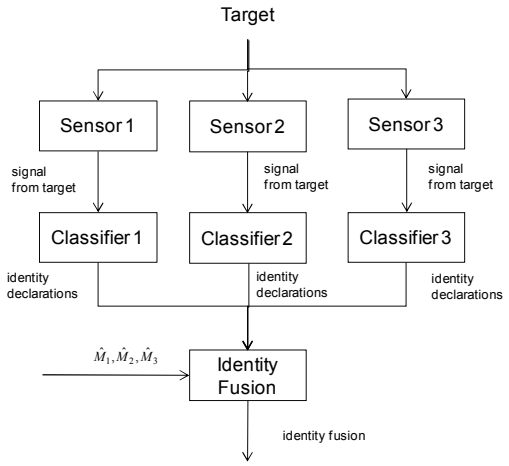


Fig. 1. Block diagram of the fusion system.

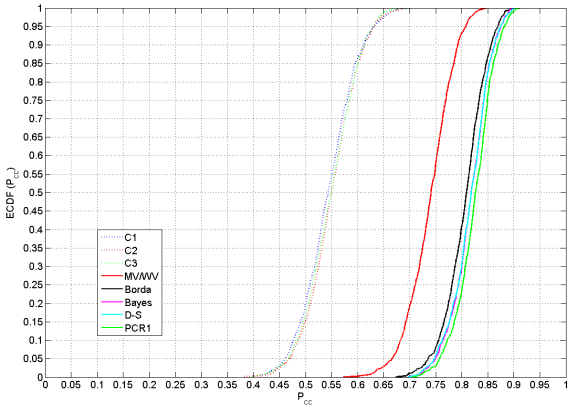


Fig. 2. ECDF versus Pcc with complementary confusion matrices (estimation from 30 samples per class).

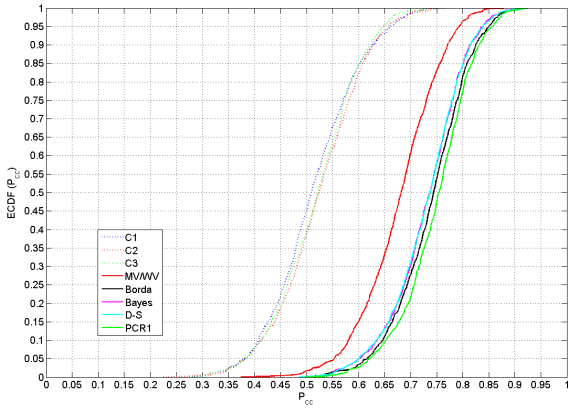


Fig. 3. ECDF versus Pcc with complementary confusion matrices (estimation from 10 samples per class).

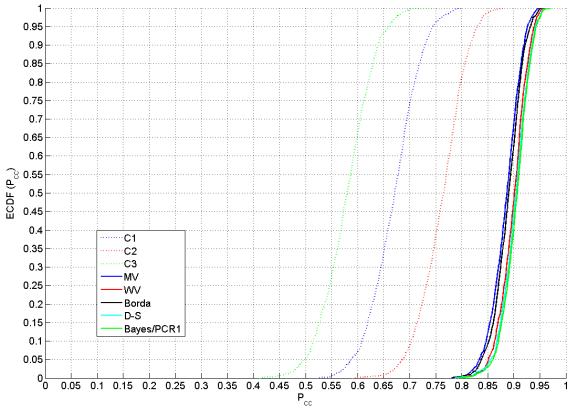


Fig. 4. ECDF versus Pcc with supplementary confusion matrices (estimation from 30 samples per class).

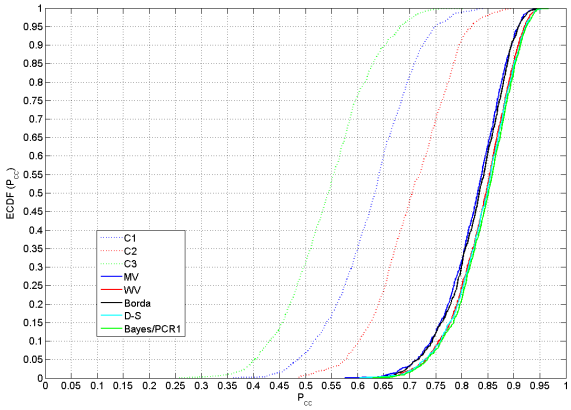


Fig. 5. ECDF versus Pcc with supplementary confusion matrices (estimation from 10 samples per class).

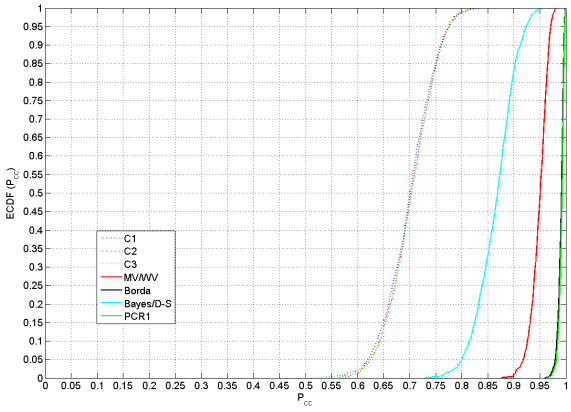


Fig. 6. ECDF versus Pcc with complementary conflicting confusion matrices (estimation from 30 samples per class).

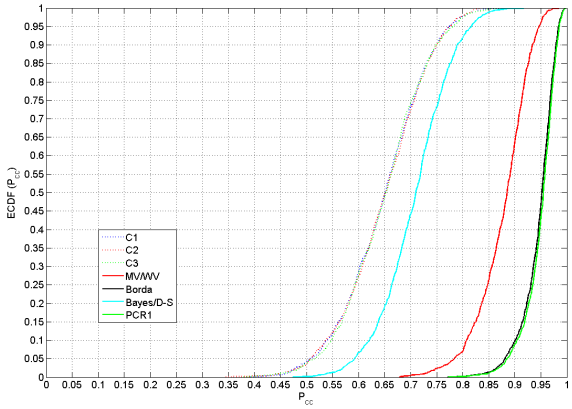


Fig. 7. ECDF versus Pcc with complementary conflicting confusion matrices (estimation from 10 samples per class).

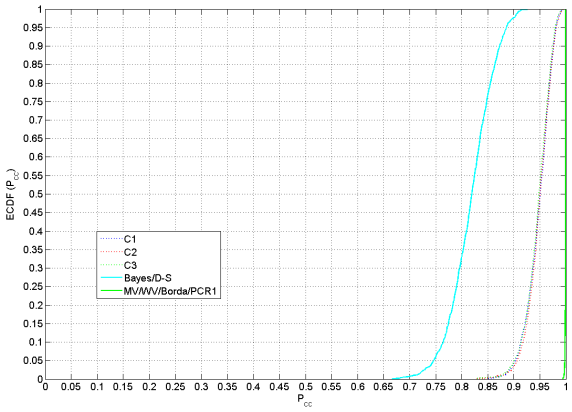


Fig. 8. ECDF of Pcc with supplementary conflicting confusion matrixes (estimation from 30 samples per class).

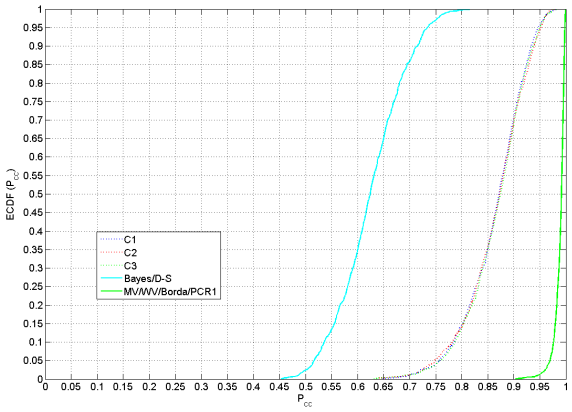


Fig. 9. ECDF of Pcc with supplementary conflicting confusion matrixes (estimation from 10 samples per class).

TABLE I. SUMMARY RESULTS (N IS THE TOTAL NUMBER OF SAMPLES).

Study cases	PROBABILITY OF CORRECT CLASSIFICATION (mean value, %)																	
	MAJORITY VOTING			WEIGHTED VOTING			BORDA COUNT			BAYES			DEMPSTER SHAFER			PCR1		
	N=30	N=60	N=90	N=30	N=60	N=90	N=30	N=60	N=90	N=30	N=60	N=90	N=30	N=60	N=90	N=30	N=60	N=90
Compl. CMs	67.8	72.4	73.9	67.8	72.4	73.9	73.9	78.9	80.7	73.1	79.3	81.6	73.0	79.4	81.6	74.8	80.5	82.4
Supp. CMs	82.3	86.9	88.4	83.9	88.4	89.9	82.7	87.4	88.8	84.3	88.9	90.4	84.0	88.7	90.2	84.3	88.9	90.4
Compl. confl. CMs	87.7	92.9	94.8	87.7	92.9	94.8	94.5	98.1	99.0	70.7	81.5	86.5	70.7	81.5	86.5	94.8	98.3	99.2
Supp. confl. CMs	98.8	99.8	99.9	98.8	99.8	99.9	98.8	99.8	99.9	62.5	75.5	81.8	62.5	75.5	81.8	98.8	99.8	99.9

Reliability and Importance Discounting of Neutrosophic Masses

Florentin Smarandache

Originally published in Smarandache, F. - *Neutrosophic Theory and its Applications, Collected Papers,* Vol. I. 2014. <hal-01092887v2>, and reprinted with permission.

! ÂÓÔÒÁÃÔ In this paper, we introduce for the first time the discounting of a neutrosophic mass in terms of reliability and respectively the importance of the source.

We show that reliability and importance discounts commute when dealing with classical masses.

ρ) Î ÔÒÏÄÕÃÔÉÏ Î Let $\Phi = \{\Phi_1, \Phi_2, \dots, \Phi_n\}$ be the frame of discernment, where $n \geq 2$, and the set of **ÆÏÃÁÌ eleme Î ÔÓ**:

$$F = \{A_1, A_2, \dots, A_m\}, \text{for } m \geq 1, F \subset G^{\Phi}. \quad (1)$$

Let $G^{\Phi} = (\Phi, \cup, \cap, \mathcal{C})$ be the **ÆÕÓÉ n ÓÐÁÃÅ**.

A **Î ÅÕÔÒÏÓÏÐÈÉÃ Í ÁÓÓ** is defined as follows:

$$m_n: G \to [0,1]^3$$

for any $x \in G, m_n(x) = (t(x), i(x), f(x)), \quad (2)$

where $t(x) =$ believe that x will occur (truth);

$i(x) =$ indeterminacy about occurence;

and $f(x) =$ believe that x will not occur (falsity).

Simply, we say in neutrosophic logic:

$t(x) =$ believe in x;

$i(x) =$ believe in neut(x)
[the neutral of x, i.e. neither x nor anti(x)];

and $f(x) =$ believe in anti(x) [the opposite of x].

Of course, $t(x),\ i(x),\ f(x) \in [0,1]$, and

$$\sum_{x\in G}[t(x)+i(x)+f(x)] = 1, (3)$$

while

$$m_n(\phi) = (0,0,0). (4)$$

It is possible that according to some parameters (or data) a source is able to predict the believe in a hypothesis x to occur, while according to other parameters (or other data) the same source may be able to find the believe in x not occuring, and upon a third category of parameters (or data) the source may find some indeterminacy (ambiguity) about hypothesis occurence.

An element $x \in G$ is called **ÆÏÃÁÌ** if

$$n_m(x) \neq (0,0,0), (5)$$

i.e. $t(x) > 0$ or $i(x) > 0$ or $f(x) > 0$.

Any **ÃÌÁÓÓÉÃÁl mÁÓÓ**:

$$m : G^{\phi} \to [0,1] (6)$$

can be simply written as a neutrosophic mass as:

$$m(A) = (m(A), 0, 0). (7)$$

ç8 $ÉÓÃÏÕÎÔÉÎÇ Á NeuÔÒÏÏÐÈÉÃ - ÁÓÓ due ÔÏ RelÉÁÂÉÌÉÔÙ of ÔÈÅ 3ÏÕÒÃÅ8

Let $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ be the reliability coefficient of the source, $\alpha \in [0,1]^3$.

Then, for any $x \in G^\theta \setminus \{\theta, I_t\}$,

where $\theta =$ the empty set

and $I_t =$ total ignorance,

$$m_n(x)_a = \big(\alpha_1 t(x), \alpha_2 i(x), \alpha_3 f(x)\big), \quad (8)$$

and

$$m_n(I_t)_\alpha = \Bigg(t(I_t) + (1-\alpha_1) \sum_{x \in G^\theta \setminus \{\phi, I_t\}} t(x),$$

$$i(I_t) + (1-\alpha_2) \sum_{x \in G^\theta \setminus \{\phi, I_t\}} i(x), f(I_t) + (1-\alpha_3) \sum_{x \in G^\theta \setminus \{\phi, I_t\}} f(x) \Bigg) \quad (9),$$

and, of course,

$$m_n(\phi)_\alpha = (0,0,0).$$

The missing mass of each element x, for $x \neq \phi, x \neq I_t$, is transferred to the mass of the total ignorance in the following way:

$t(x) - \alpha_1 t(x) = (1-\alpha_1) \cdot t(x)$ is transferred to $t(I_t)$, (10)

$i(x) - \alpha_2 i(x) = (1-\alpha_2) \cdot i(x)$ is transferred to $i(I_t)$, (11)

and $f(x) - \alpha_3 f(x) = (1-\alpha_3) \cdot f(x)$ is transferred to $f(I_t)$. (12)

σ8 $ÉÓÃÏÕÎÔÉÎÇ Á .ÅÕÔÒÏÓÏÐÈÉÃ -ÁÓs ÄÕe ôÏ ÔÈÅ)Í ÐÏÒÔÁÎce Ïf ÔÈÅ 3ÏÕÒÃÅ8

Let $\beta \in [0,1]$ be the importance coefficient of the source. This discounting can be done in several ways.

a. For any $x \in G^\theta \setminus \{\phi\}$,

$$m_n(x)_{\beta_1} = \big(\beta \cdot t(x), i(x), f(x) + (1-\beta) \cdot t(x)\big), (13)$$

which means that $t(x)$, the believe in x, is diminished to $\beta \cdot t(x)$, and the missing mass, $t(x) - \beta \cdot t(x) = (1-\beta) \cdot t(x)$, is transferred to the believe in $anti(x)$.

b. Another way:

For any $x \in G^\theta \setminus \{\phi\}$,

$$m_n(x)_{\beta_2} = \big(\beta \cdot t(x), i(x) + (1-\beta) \cdot t(x), f(x)\big), (14)$$

which means that $t(x)$, the believe in x, is similarly diminished to $\beta \cdot t(x)$, and the missing mass $(1-\beta) \cdot t(x)$ is now transferred to the believe in $neut(x)$.

c. The third way is the most general, putting together the first and second ways.

For any $x \in G^\theta \setminus \{\phi\}$,

$$m_n(x)_{\beta_3} = \big(\beta \cdot t(x), i(x) + (1-\beta) \cdot t(x) \cdot \gamma, f(x) + (1-\beta) \cdot t(x) \cdot (1-\gamma)\big), (15)$$

where $\gamma \in [0,1]$ is a parameter that splits the missing mass $(1-\beta) \cdot t(x)$ a part to $i(x)$ and the other part to $f(x)$.

For $\gamma = 0$, one gets the first way of distribution, and when $\gamma = 1$, one gets the second way of distribution.

τ8 $ÉÓÃÏÕÎÔÉÎÇ Ïf 2ÀÌÉÁÂÉÌÉôy ÁÎÄ) Í ÐÏòtance Ïf 3ÏÕÒces én 'ÅÎÅÒÁÌ $Ï Not Co Í Í ÕÔÅ8

a. Reliability first, Importance second.

For any $x \in G^\theta \setminus \{\phi, I_t\}$, one has after reliability α discounting, where

$$\alpha = (\alpha_1, \alpha_2, \alpha_3):$$

$$m_n(x)_\alpha = \big(\alpha_1 \cdot t(x), \alpha_2 \cdot t(x), \alpha_3 \cdot f(x)\big), \quad (16)$$

and
$$m_n(I_t)_\alpha = \Bigg(t(I_t) + (1-\alpha_1) \cdot \sum_{x \in G^\theta \setminus \{\phi, I_t\}} t(x)\,, i(I_t) + (1-\alpha_2) \cdot \sum_{x \in G^\theta \setminus \{\phi, I_t\}} i(x)\,, f(I_t) + (1-\alpha_3) \cdot \sum_{x \in G^\theta \setminus \{\phi, I_t\}} f(x) \Bigg)$$
$$\stackrel{\text{def}}{=} \big(T_{I_t}, I_{I_t}, F_{I_t}\big). \quad (17)$$

Now we do the importance β discounting method, the third importance discounting way which is the most general:

$$m_n(x)_{\alpha\beta_3} = \big(\beta\alpha_1 t(x), \alpha_2 i(x) + (1-\beta)\alpha_1 t(x)\gamma, \alpha_3 f(x) + (1-\beta)\alpha_1 t(x)(1-\gamma)\big) \quad (18)$$

and

$$m_n(I_t)_{\alpha\beta_3} = \Big(\beta \cdot T_{I_t}, I_{I_t} + (1-\beta)T_{I_t} \cdot \gamma, F_{I_t} + (1-\beta)T_{I_t}(1-\gamma)\Big). \quad (19)$$

b. Importance first, Reliability second.

For any $x \in G^\theta \setminus \{\phi, I_t\}$, one has after importance β discounting (third way):

$$m_n(x)_{\beta_3} = \big(\beta \cdot t(x), i(x) + (1-\beta)t(x)\gamma, f(x) + (1-\beta)t(x)(1-\gamma)\big) \quad (20)$$

and

$$m_n(I_t)_{\beta_3} = \Big(\beta \cdot t\big(I_t\big), i(I_{I_t}) + (1-\beta)t(I_t)\gamma,\ f(I_t) + (1-\beta)t(I_t)(1-\gamma)\Big). \quad (21)$$

Now we do the reliability $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ discounting, and one gets:

$$m_n(x)_{\beta_3\alpha} = \big(\alpha_1 \cdot \beta \cdot t(x), \alpha_2 \cdot i(x) + \alpha_2(1-\beta)t(x)\gamma, \alpha_3 \cdot f(x) + \alpha_3 \cdot (1-\beta)t(x)(1-\gamma)\big) \quad (22)$$

and

$$m_n(I_t)_{\beta_3\alpha} = \big(\alpha_1 \cdot \beta \cdot t(I_t), \alpha_2 \cdot i(I_t) + \alpha_2(1-\beta)t(I_t)\gamma, \alpha_3 \cdot f(I_t) + \alpha_3(1-\beta)t(I_t)(1-\gamma)\big). \quad (23)$$

2emark.

We see that (a) and (b) are in general different, so reliability of sources does not commute with the importance of sources.

υ8 0ÁÒÔÉÃÕÌÁr #Áse ×Èen 2ÅÌÉÁÂÉÌÉty ÁÎd) Í ÐÏÒÔÁÎÃÅ $ÉÓÃÏÕÎÔÉg ÏÆ - ÁÓÓÅÓ Co Í Í ÕÔÅ8

Let's consider a classical mass $m: G^\theta \to [0,1]$ (24)

and the focal set $F \subset G^\theta$, $F = \{A_1, A_2, \dots, A_m\}, m \geq 1$, (25)

and of course $m(A_i) > 0$, for $1 \leq i \leq m$.

Suppose $m(A_i) = a_i \in (0,1]$. (26)

a. Reliability first, Importance second.

Let $\alpha \in [0,1]$ be the reliability coefficient of $m(\cdot)$.

For $x \in G^\theta \setminus \{\phi, I_t\}$, one has $m(x)_\alpha = \alpha \cdot m(x)$, (27)

and $m(I_t) = \alpha \cdot m(I_t) + 1 - \alpha$. (28)

Let $\beta \in [0,1]$ be the importance coefficient of $m(\cdot)$.

Then, for $x \in G^\theta \setminus \{\phi, I_t\}$,

$$m(x)_{\alpha\beta} = \big(\beta\alpha m(x), \alpha m(x) - \beta\alpha m(x)\big) = \alpha \cdot m(x) \cdot (\beta, 1-\beta), \quad (29)$$

considering only two components: believe that x occurs and, respectively, believe that x does not occur.

Further on,

$$m(I_t)_{\alpha\beta} = (\beta\alpha m(I_t) + \beta - \beta\alpha, \alpha m(I_t) + 1 - \alpha - \beta\alpha m(I_t) - \beta + \beta\alpha) = [\alpha m(I_t) + 1 - \alpha] \cdot (\beta, 1-\beta). \quad (30)$$

b. Importance first, Reliability second.

For $x \in G^{\theta} \setminus \{\phi, I_t\}$, one has

$$m(x)_{\beta} = \big(\beta \cdot m(x), m(x) - \beta \cdot m(x)\big) = m(x) \cdot (\beta, 1-\beta), \quad (31)$$

$$\text{and } m(I_t)_{\beta} = \big(\beta m(I_t), m(I_t) - \beta m(I_t)\big) = m(I_t) \cdot (\beta, 1-\beta). \quad (32)$$

Then, for the reliability discounting scaler α one has:

$$m(x)_{\beta\alpha} = \alpha m(x)(\beta, 1-\beta) = \big(\alpha m(x)\beta, \alpha m(x) - \alpha\beta m(m)\big) \quad (33)$$

$$\text{and } m(I_t)_{\beta\alpha} = \alpha \cdot m(I_t)(\beta, 1-\beta) + (1-\alpha)(\beta, 1-\beta) = [\alpha m(I_t) + 1 - \alpha] \cdot (\beta, 1-\beta) = (\alpha m(I_t)\beta, \alpha m(I_t) - \alpha m(I_t)\beta) + (\beta - \alpha\beta, 1 - \alpha - \beta + \alpha\beta) = (\alpha\beta m(I_t) + \beta - \alpha\beta, \alpha m(I_t) - \alpha\beta m(I_t) + 1 - \alpha - \beta - \alpha\beta). \quad (34)$$

Hence (a) and (b) are equal in this case.

φ§ %ØÁ Í ÐÌÅÓ§

1. Classical mass.

The following classical is given on $\theta = \{A, B\}$:

	A	B	AUB
m	0.4	0.5	0.1

(35)

Let $\alpha = 0.8$ be the reliability coefficient and $\beta = 0.7$ be the importance coefficient.

a. Reliability first, Importance second.

	A	B	AUB
m_{α}	0.32	0.40	0.28
$m_{\alpha\beta}$	(0.224, 0.096)	(0.280, 0.120)	(0.196, 0.084)

(36)

We have computed in the following way:

$$m_\alpha(A) = 0.8m(A) = 0.8(0.4) = 0.32, (37)$$

$$m_\alpha(B) = 0.8m(B) = 0.8(0.5) = 0.40, (38)$$

$$m_\alpha(AUB) = 0.8(\text{AUB}) + 1 - 0.8 = 0.8(0.1) + 0.2 = 0.28, (39)$$

and $m_{\alpha\beta}(B) = \big(0.7m_\alpha(A), m_\alpha(A) - 0.7m_\alpha(A)\big) = \big(0.7(0.32), 0.32 - 0.7(0.32)\big) = (0.224, 0.096), (40)$

$$m_{\alpha\beta}(B) = \big(0.7m_\alpha(B), m_\alpha(B) - 0.7m_\alpha(B)\big) = \big(0.7(0.40), 0.40 - 0.7(0.40)\big) = (0.280, 0.120), (41)$$

$$m_{\alpha\beta}(AUB) = \big(0.7m_\alpha(AUB), m_\alpha(AUB) - 0.7m_\alpha(AUB)\big) = \big(0.7(0.28), 0.28 - 0.7(0.28)\big) = (0.196, 0.084). (42)$$

b. Importance first, Reliability second.

	A	B	AUB
m	0.4	0.5	0.1
m_β	(0.28, 0.12)	(0.35, 0.15)	(0.07, 0.03)
$m_{\beta\alpha}$	(0.224, 0.096	(0.280, 0.120)	(0.196, 0.084)

(43)

We computed in the following way:

$$m_\beta(A) = \big(\beta m(A), (1-\beta)m(A)\big) = \big(0.7(0.4), (1-0.7)(0.4)\big) = (0.280, 0.120), (44)$$

$$m_\beta(B) = \big(\beta m(B), (1-\beta)m(B)\big) = \big(0.7(0.5), (1-0.7)(0.5)\big) = (0.35, 0.15), (45)$$

$$m_\beta(AUB) = \big(\beta m(AUB), (1-\beta)m(AUB)\big) = \big(0.7(0.1), (1-0.1)(0.1)\big) = (0.07, 0.03), (46)$$

and $m_{\beta\alpha}(A) = \alpha m_\beta(A) = 0.8(0.28, 0.12) = \big(0.8(0.28), 0.8(0.12)\big) = (0.224, 0.096), (47)$

$$m_{\beta\alpha}(B) = \alpha m_\beta(B) = 0.8(0.35, 0.15) = \big(0.8(0.35), 0.8(0.15)\big) = (0.280, 0.120), (48)$$

$$m_{\beta\alpha}(AUB) = \alpha m(AUB)(\beta, 1-\beta) + (1-\alpha)(\beta, 1-\beta) = 0.8(0.1)(0.7, 1-0.7) + (1-0.8)(0.7, 1-0.7) = 0.08(0.7, 0.3) + 0.2(0.7, 0.3) = (0.056, 0.024) + (0.140, 0.060) = (0.056 + 0.140, 0.024 + 0.060) = (0.196, 0.084). \quad (49)$$

Therefore reliability discount commutes with importance discount of sources when one has classical masses.

The result is interpreted this way: believe in A is 0.224 and believe in $nonA$ is 0.096, believe in B is 0.280 and believe in $nonB$ is 0.120, and believe in total ignorance AUB is 0.196, and believe in non-ignorance is 0.084.

χ8 3Á Í Å %ØÁ Í ÐÌÅ ×ÉÔh $ÉÆÆÅÒÅÎÔ 2ÅÄÉÓÔÒÉÂÕÔÉÏ Î Ïf - ÁÓÓÅÓ 2ÅÌÁÔÅÄ ÔÏ) Í ÐÏÒÔÁ Î ÃÅ of 3ÏÕÒÃÅÓ8

Let's consider the third way of redistribution of masses related to importance coefficient of sources. $\beta = 0.7$, but $\gamma = 0.4$, which means that 40% of β is redistributed to $i(x)$ and 60% of β is redistributed to $f(x)$ for each $x \in G^{\theta} \setminus \{\phi\}$; and $\alpha = 0.8$.

a. Reliability first, Importance second.

	A	B	AUB
m	0.4	0.5	0.1
m_{α}	0.32	0.40	0.28
$m_{\alpha\beta}$	(0.2240, 0.0384, 0.0576)	(0.2800, 0.0480, 0.0720)	(0.1960, 0.0336, 0.0504).

(50)

We computed m_α in the same way.

But:

$$m_{\alpha\beta}(A) = \big(\beta \cdot m_\alpha(A), i_\alpha(A) + (1-\beta)m_\alpha(A)\cdot\gamma, f_\alpha(A) + (1-\beta)m_\alpha(A)(1-\gamma)\big) = \big(0.7(0.32), 0 + (1-0.7)(0.32)(0.4), 0 + (1-0.7)(0.32)(1-0.4)\big) = (0.2240, 0.0384, 0.0576). \quad (51)$$

Similarly for $m_{\alpha\beta}(B)$ and $m_{\alpha\beta}(AUB)$.

b. Importance first, Reliability second.

	A	B	AUB
m	0.4	0.5	0.1
m_β	(0.280, 0.048, 0.072)	(0.350, 0.060, 0.090)	(0.070, 0.012, 0.018)
$m_\beta\alpha$	(0.2240, 0.0384, 0.0576)	(0.2800, 0.0480, 0.0720)	(0.1960, 0.0336, 0.0504).

(52)

We computed $m_\beta(\cdot)$ in the following way:

$$m_\beta(A) = \big(\beta \cdot t(A), i(A) + (1-\beta)t(A)\cdot\gamma, f(A) + (1-\beta)t(A)(1-\gamma)\big) = \big(0.7(0.4), 0 + (1-0.7)(0.4)(0.4), 0 + (1-0.7)0.4(1-0.4)\big) = (0.280, 0.048, 0.072). \quad (53)$$

Similarly for $m_\beta(B)$ and $m_\beta(AUB)$.

To compute $m_{\beta\alpha}(\cdot)$, we take $\alpha_1 = \alpha_2 = \alpha_3 = 0.8$, (54)

in formulas (8) and (9).

$$m_{\beta\alpha}(A) = \alpha \cdot m_{\beta}(A) = 0.8(0.280, 0.048, 0.072)$$
$$= \big(0.8(0.280), 0.8(0.048), 0.8(0.072)\big)$$
$$= (0.2240, 0.0384, 0.0576). (55)$$

Similarly $m_{\beta\alpha}(B) = 0.8(0.350, 0.060, 0.090) =$
$(0.2800, 0.0480, 0.0720). (56)$

For $m_{\beta\alpha}(AUB)$ we use formula (9):

$$m_{\beta\alpha}(AUB) = \big(t_{\beta}(AUB) + (1-\alpha)\big[t_{\beta}(A) + t_{\beta}(B)\big],\ i_{\beta}(AUB)$$
$$+ (1-\alpha)\big[i_{\beta}(A) + i_{\beta}(B)\big],$$
$$f_{\beta}(AUB) + (1-\alpha)\big[f_{\beta}(A) + f_{\beta}(B)\big]\big)$$
$$= (0.070 + (1-0.8)[0.280 + 0.350], 0.012$$
$$+ (1-0.8)[0.048 + 0.060], 0.018 + (1-0.8)[0.072 + 0.090])$$
$$= (0.1960, 0.0336, 0.0504).$$

Again, the reliability discount and importance discount commute.

ψ8 #Ï ÎÃÌÕÓÉÏ Î8

In this paper we have defined a new way of discounting a classical and neutrosophic mass with respect to its importance. We have also defined the discounting of a neutrosophic source with respect to its reliability.

In general, the reliability discount and importance discount do not commute. But if one uses classical masses, they commute (as in Examples 1 and 2).

!ÃËÎÏ×ÌÅÄÇÅÍÅÎô

The author would like to thank Dr. Jean Dezert for his opinions about this paper.

2Å£ÅÒÅÎÃÅó

1. F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of Sources of Evidence with Different Importances and Reliabilities*, Fusion 2010 International Conference, Edinburgh, Scotland, 26-29 July, 2010.

2. Florentin Smarandache, *Neutrosophic Masses & Indeterminate Models. Applications to Information Fusion*, Proceedings of the International Conference on Advanced Mechatronic Systems [ICAMechS 2012], Tokyo, Japan, 18-21 September 2012.

Part 2:

Applications of DSmT

DSm Theory for Fusing Highly Conflicting ESM Reports

Pierre Valin
Pascal Djiknavorian
Dominic Grenier

Originally published as Valin P., Djiknavorian P., Grenier D., *DSm theory for fusing highly conflicting ESM reports*, in Proc. of Fusion 2009, Seattle, WA, USA, 6-9 July 2009, and reprinted with permission.

Abstract - *Electronic Support Measures consist of passive receivers which can identify emitters coming from a small bearing angle, which, in turn, can be related to platforms that belong to 3 classes: either Friend, Neutral, or Hostile. Decision makers prefer results presented in STANAG 1241 allegiance form, which adds 2 new classes: Assumed Friend, and Suspect. Dezert-Smarandache (DSm) theory is particularly suited to this problem, since it allows for intersections between the original 3 classes. Results are presented showing that the theory can be successfully applied to the problem of associating ESM reports to established tracks, and its results identify when miss-associations have occurred and to what extent. Results are also compared to Dempster-Shafer theory which can only reason on the original 3 classes. Thus decision makers are offered STANAG 1241 allegiance results in a timely manner, with quick allegiance change when appropriate and stability in allegiance declaration otherwise.*

Keywords: Electronic Support Measures, Dezert-Smarandache, Dempster-Shafer, allegiance, fusion.

1 Introduction

Electronic Support Measures (ESM) consist of passive receivers which can identify emitters coming from a small bearing angle, but cannot determine range (although some are in development to provide a rough measure of range). The detected emitters can be related to platforms that belong to 3 classes: either Friend (F=1), Neutral (N=2) or Hostile (H=3), heretofore called ESM-allegiance, within that bearing angle.

In the case of dense targets, ESM-allegiance can fluctuate wildly due to miss-associations of an ESM report to established track. Hence, decision makers would like the target platforms to be identified on a more refined basis, belonging to 5 classes: Hostile (or Foe), Suspect (S), Neutral, Assumed Friend (AF), and Friend, since they realize that no fusion algorithm can be perfect and would prefer some stability in an allegiance declaration, rather than oscillations between extremes. This will heretofore be referred to as STANAG 1241 allegiance, or just STANAG allegiance for short [1].

With this more refined STANAG-allegiance, a decision maker would probably take no aggressive action against either a friend or an assumed friend (although he/she would monitor an assumed friend more closely). Similarly a decision maker would probably take aggressive action against a foe and send a reconnaissance force (or a warning salvo) towards a suspect. Neutral platforms would correspond to countries not involved in the current conflict.

All incoming sensor declarations correspond to a frame of discernment of 3 classes, and several theories exist to treat a series of such declarations to obtain a fused result in the same frame of discernment, like Bayesian reasoning and Dempster-Shafer (DS) reasoning [2, 3] (often called evidence theory). However, when the output frame of discernment is larger that the input frame of discernment, an interpretation has to be made as to what this could mean, or how that could be generated. This is the subject of the next section.

1.1 Some solutions

It should be noted that Bayes theory is implemented in a very complex form in STANAG 4162 [4], and that DS theory is found on board many platforms, such as the German F124 frigates [5], the Finnish Fast Attack Craft Squadron 2000 [6], and the Light Airborne Multi-Purpose System (LAMPS) helicopters of the US Navy [7]. The translation from DS to Bayes can be performed via the pignistic transformation [8], and the result broadcast via tactical data links.

In all these implementations, the emitter detected is first correlated to a platform, and then to an allegiance. According to [9], recognition of a platform can range from a very rough scale (e.g. combatant/merchant) to a very fine one (e.g. name of contact/track), whereas identification refers to the assignment of one of the 6 standard STANAG 1241 identities (for which we adopt the word “allegiance” in this report) to a track. The extra identity is “unknown”, which we disregard in this report, assuming that all detected emitters are identifiable.

Therefore, this report investigates an alternative method of achieving STANAG-allegiances, which does not aim to compete with the above implementations, but rather can be seen as an expert advisor to the decision maker. Since Dezert-Smarandache theory was only developed extensively after the publications of the STANAGs, this could not have been foreseen by NATO, and is thus worthy of experimentation.

1.2 An interpretation of STANAG 1241

Dezert-Smarandache (DSm) theory can coherently, with well-defined fusion rules, lead to an output amongst 5 classes, even though the input classes number only 3, because the theory allows for intersections. For example,

- "Suspect" might be the result obtained after fusing "Hostile" with "Neutral", and
- "Assumed Friend" might be the result obtained after fusing "Friend' with "Neutral".

This illustrated in the Venn diagram of Figure 1 below.

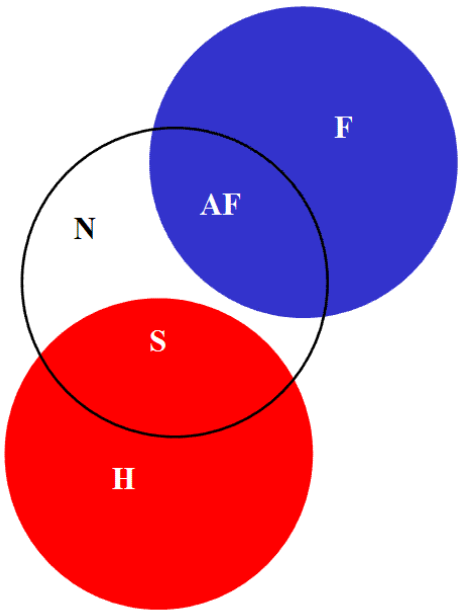


Figure 1. Venn diagram for the STANAG allegiances.

Note that the set intersection $1 \cap 3 = \emptyset$, the null set, which is a constraint in DSm, leading to the use of its hybrid rule. It also corresponds to the most likely mission for Canadian Forces (CF), namely peace-keeping, or general surveillance, when hostile and friendly forces are not likely to be located close to each other.

1.3 Another interpretation of STANAG 1241

The interpretation in the preceding sub-section is a conservative one, namely that there is only one easy way to become suspect. This could correspond to a decision maker being in a non-threatening situation due to the choice of mission, e.g. peace-keeping. There could be situations where there is a need for a more aggressive response. In the case of a combat mission for example, the appropriate Venn diagram might be the one of Figure 2, where there are many more ways to become suspect, namely all the intersections bordering Hostile.

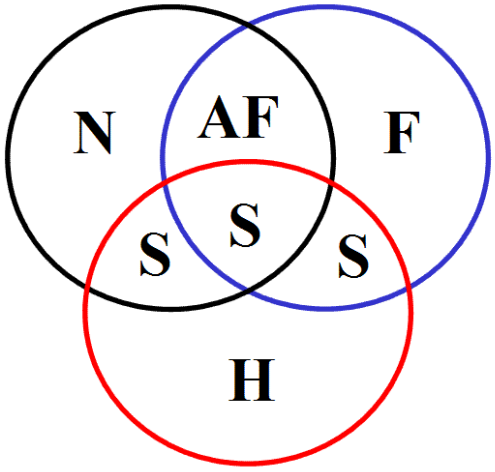


Figure 2. Another possible Venn diagram.

Figure 2 corresponds to a combat situation more appropriate for the USA, or to the CF as long as they play an active role in the Kandahar region of Afghanistan. The situation of Figure 1 will be the one implemented in this paper, as it is more in line with CF roles, and also because all of the features of DSm theory can be exercised, without the additional complexity of keeping all the intersections of Figure 1.

2 Dezert-Smarandache Theory

2.1 Formulae for DS and DSm theories

Since DS theory has been in use for over 40 years, the reader is assumed to be familiar with it [2, 3]. DSm theory encompasses DS theory as a special case, namely when all intersections are null. Both use the language of masses assigned to each declaration from a sensor (in our case, the ESM sensor). A declaration is a set made up of singletons of the frame of discernment Θ, and all sets that can be made from them through unions are allowed (this is referred to as the power set 2^Θ of DS theory). In DSm theory, all unions and intersections are allowed for a declaration, this forming the much larger hyper power set D^Θ. For our special case of cardinality 3, $\Theta = \{\theta_1, \theta_2, \theta_3\}$, with $|\Theta| = 3$, D^Θ is still of manageable size, namely has a cardinality of 19.

In DST, a combined "fused" mass is obtained by combining the previous (presumably the results of previous fusion steps) $m_1(A)$ with a new $m_2(B)$ to obtain a new fused result by applying the conjunctive rule

$$m_1 \oplus m_2(C) = \Sigma\, m_1(A)\, m_2(B) \qquad (1)$$

where $C = A \cap B$, and by re-normalizing by $(1-K)^{-1}$ where K is the conflict corresponding to the sum of all masses for which the set intersection yields the null set. This common renormalization is a critical feature of DS theory, and allows for it to be associative, whereas a multitude of

alternate ways of redistributing the conflict (proposed by numerous authors) loses this property. The associativity of DST is key when the time tags of the sensor reports are unreliable, since associative rules are impervious to a different order of reports coming in, but all others rules can be extremely sensitive to the order of reports. This is the main reason we concentrate only on DS vs. DSm, but another reason is the proliferation of alternatives to DS, which redistribute the conflict in various fashions (for a review, see [10]).

In DSm theory, a constraint like the one that was imposed by Figure 1, namely that $1 \cap 3 = \emptyset$ is treated by the hybrid DSm rule below:

$$m(A) = \phi(A) [S_1(A) + S_2(A) + S_3(A)] \qquad (2)$$

The reader is referred to a series of books [10, 11] on DSm theory for lengthy descriptions of the meaning of this formula (note that the function ϕ is not to be confused with the empty set). A three-step approach is proposed in chapter 5 of [11], which is used here.

If the incoming sensor reports are in DS-space: Friend (F=1), Neutral (N=2) or Hostile (H=3), then Figure 1 has the interpretation in DSm space (allowing intersections during the fusion step) of:

$$\text{Friend} = \{\theta_1 - \theta_1 \cap \theta_2\}$$

$$\text{Hostile} = \{\theta_3 - \theta_3 \cap \theta_2\}$$

$$\text{Assumed Friend} = \{\theta_1 \cap \theta_2\}$$

$$\text{Suspect} = \{\theta_2 \cap \theta_3\}$$

$$\text{Neutral} = \{\theta_2 - \theta_1 \cap \theta_2 - \theta_3 \cap \theta_2\}$$

As expected, all STANAG-allegiances (masses assigned to the sets mentioned above) sum up to 1, as shown below. The left hand side, which is the sum of the masses for all 5 classes, yields the right hand side, which is unity in DSm theory.

$$\theta_1 - \theta_1 \cap \theta_2 + \theta_3 - \theta_3 \cap \theta_2 + \theta_1 \cap \theta_2 + \theta_2 \cap \theta_3 + \theta_2 - \theta_1 \cap \theta_2 - \theta_3 \cap \theta_2 = \theta_1 + \theta_2 + \theta_3 - \theta_1 \cap \theta_2 - \theta_3 \cap \theta_2 = 1 \qquad (3)$$

(since $m(\theta_1 \cap \theta_3) = 0$, i.e. $\theta_1 \cap \theta_3 = 1 \cap 3 = \emptyset$ by Figure 1).

2.2 A typical simulation scenario

In order to compare DS with DSm, one must list the pre-requisites that the scenario must address. It must:

- be able to adequately represent the known ground truth
- contain sufficient miss-associations to be realistic and to test the robustness of the theories
- only provide partial knowledge about the ESM sensor declaration, which therefore contains uncertainty
- be able to show stability under countermeasures, yet
- be able to switch allegiance when the ground truth does so

The following scenario parameters have therefore been chosen accordingly:

- Ground truth is FRIEND for the first 50 iterations of the scenario and HOSTILE for the last 50.
- the number of correct associations is 80%, corresponding to countermeasures appearing 20% of the time, in a randomly selected sequence.
- the ESM declaration has a mass (confidence value in Bayesian terms) of 0.7, with the rest (0.3) being assigned to the ignorance (the full set of elements, namely Θ).

The last 2 bullets of the first list would translate into stability for the first 50 iterations and eventual stability for the last 50 iterations, after the allegiance switch at iteration 50.

This scenario will be the one addressed in the next section, while a Monte-Carlo study is described in the subsequent section. Each Monte-Carlo run corresponds to a different realization using the above scenario parameters, but with a different random seed. The scenario chosen is depicted in Figure 3 below.

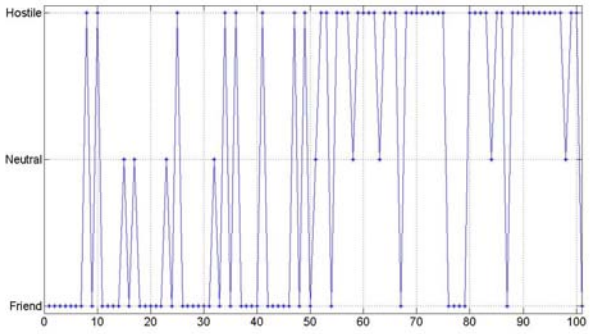


Figure 3. Chosen scenario.

The vertical axis represents the allegiance Friend, Neutral, or Hostile. Roughly 80% of the time the ESM declares the correct allegiance according to ground truth, and the remaining 20% is roughly equally split between the other two allegiances. There is an allegiance switch at the 50th iteration, and the selected randomly selected seed in the above generated scenario generates a rather unusual sequence of 4 false Friend declarations starting at iteration 76 (when actually Hostile is the ground truth), which will be very challenging for the theories.

3 Results for the simulated scenario

Before presenting the results for DS, it should be noted that the original form of DS tends to be overly optimistic. Given enough evidence concerning an allegiance, it will be very hard for it to change allegiances at iteration 50. This is a well-known problem, and a well-known ad hoc solution

exists [12], and consists in renormalizing after each fusion step by giving a value to the complete ignorance which can never be below a certain factor (chosen here to be 0.02). Comparison will be made with DSm and the Proportional Conflict Redistribution (PCR) #5 (PCR5) preferred by Dezert and Smarandache [10].

3.1 DS results

The result for DST is shown in Figure 4 below with Friend (1), Neutral (2) and Hostile (3).

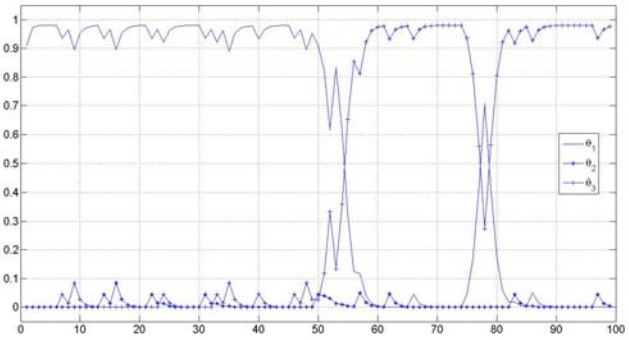


Figure 4. DS result for the chosen scenario.

DS never becomes confused, reaches the ESM-allegiance quickly and maintains it until iteration 50. It then reacts reasonably rapidly and takes about 6 reports before switching allegiance as it should. Furthermore after being confused for an iteration around the sequence of 4 Friend reports starting at iteration 76, it quickly reverts to the correct Hostile status.

Note that a decision maker could look at this curve and see an oscillation pointing to miss-associations without being able to clearly distinguish between a miss-association with the other two possible allegiances. This fairly quick reaction is due to the 0.02 assigned to the ignorance, which translates to DS never being more than 98% sure of an ESM-allegiance, as can be seen by the curve topping out at 0.98. Figure 4 shows the mass, which is also the pignistic probability for this case, with the latter being normally used to make a decision.

3.2 DSm results

For the hybrid DSm rule [10], it was suggested to use the Generalized Pignistic Probability in order to make a decision on a singleton belonging to the input ESM-allegiance. This seems to cause problems [13]. Since the whole idea behind using DSm was to present the results to the decision maker in the STANAG allegiance format, the result of Figure 5 would be shown to the decision maker.

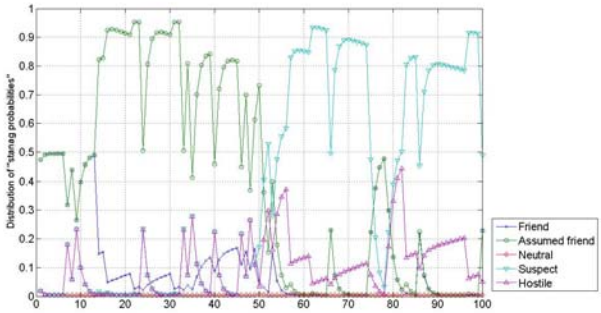


Figure 5. DSm result for the chosen scenario.

The decision maker would clearly be informed that miss-associations have occurred, since Assumed Friend dominates for the first 50 iterations and Suspect for the latter 50. DSm is more susceptible to miss-associations than DS (the dips are more pronounced), but it has the advantage of giving extra information to the decision maker, namely that the fusion algorithm is having difficulty with associating ESM reports to established tracks.

Just as for DS, the Friend declarations starting at iteration 76 cause confusion, as it should. The change in allegiance at iteration 50 is detected nearly as fast as DST. What is even more important is that F and AF are clearly preferred for the first 50 iterations and S and H for the last 50, as they should.

3.3 PCR5 results

PCR5 shows a similar behaviour, but is much less sure of what's going on (the peaks are not as pronounced), as seen in Figure 6. Again, F and AF are clearly preferred for the first 50 iterations and S and H for the last 50, as they should.

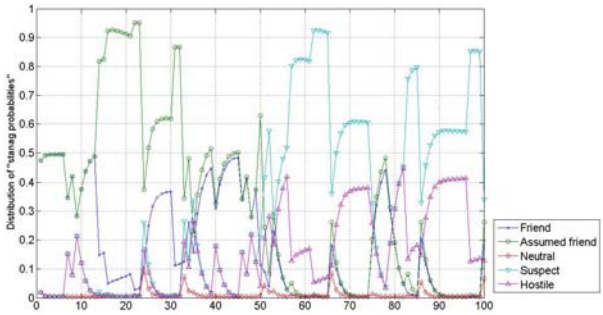


Figure 6. PCR5 result for the chosen scenario.

3.4 Decision-making threshold

Because of the occasionally oscillatory nature of some combination rules, one has to ask oneself when to make a decision or recommend one to the commander. This is illustrated in Figure 7 for DS although the same is

applicable for all the others. A threshold at a very secure 90% would result in a longer time for allegiance change, and result in a longer period of indecision around iteration 76, compared to one at 70%.

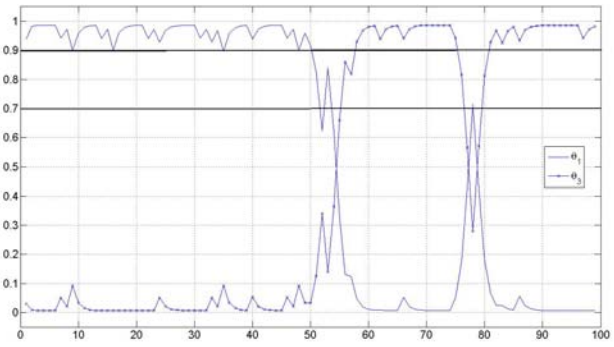


Figure 7. Decision thresholds.

4 Monte-Carlo results

Although a special case such as the one described in the previous section offers valuable insight, one might question if the conclusions from that one scenario pass the test of multiple Monte-Carlo scenarios. This question is answered in this section.

In order to sample the parameter space in a different way, the simulations below correspond to 90% correct associations (higher than the previous 80%), an ESM confidence at 60% (lower than the previous 70%) and an ignorance threshold at 0.02 as before. The number of Monte-Carlo runs was set to 100.

4.1 DS results

The result for DS is shown in Figure 8. As expected, since DS reasons over the 3 input classes, Suspect and Assumed Friend are not involved.

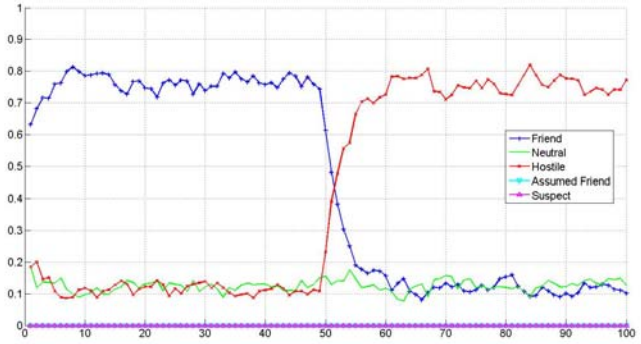


Figure 8. DS result after 100 Monte-Carlo runs.

Naturally, since Assumed Friend and Suspect do not exist in DST, these are calculated as zero. Friend, Neutral, and Hostile have the expected behaviour. One sees the same response times, after an average over 100 runs, as was seen in the selected scenario of the previous section.

4.2 DSm results

The similar result for DSm is shown in Figure 9 below. In this case, AF dominates for the first 50 iterations, on average (over 100 runs) and S for the last 50, confirming that the chosen scenario was representative of the behaviour of DSm. The response times are similar on average also. DSm is slightly less sure (plateau at 70%) than DS (plateau at 80%), but this can be adjusted by lowering the decision threshold accordingly.

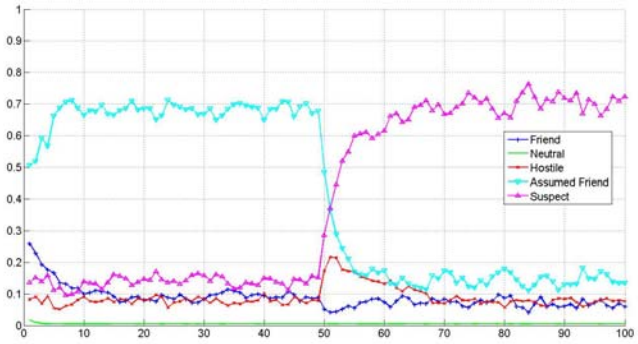


Figure 9. DSm result after 100 Monte-Carlo runs.

4.3 PCR5 results

Finally, the PCR5 result is shown in Figure 10 below. In this case also, AF dominates for the first 50 iterations, on average (over 100 runs), and S for the last 50, confirming that the chosen scenario was representative of the behaviour of PCR5. The response times are similar on average also. PCR5 is slightly less sure (plateau at 60%) than DST (plateau at 80%) or DSmT (plateau at 70%).

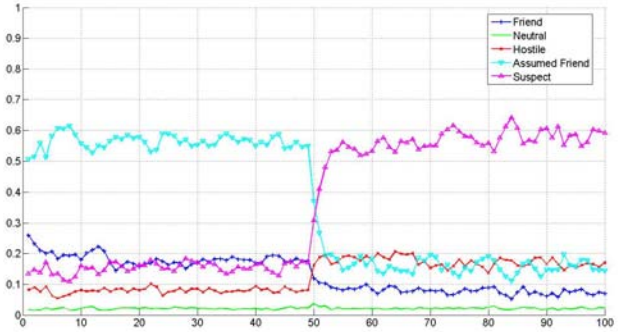


Figure 10. PCR5 result after 100 Monte-Carlo runs.

4.4 Effect of varying the ESM parameters

In order to study the effects of varying the ESM parameters, the simulations below correspond to an ESM confidence at 80% (higher than the previous 60%) and an ignorance

threshold at 0.05 (higher than the 0.02 used previously). The number of Monte-Carlo runs was again set to 100.

A filter was also applied to the input ESM declarations over a window of 4 iterations. The filter assigns lesser confidence to ESM reports which are not well represented in the window. More on this sliding window filtering is available in [13]. The idea of such a sliding window has also been studied before with good results for a variety of reasoning schemes [14]. The results are shown in Figure 11 for DS, Figure 12 for DSm and Figure 13 for PCR5. From these figures, one can see the smoothing effect of the filter, but more importantly the all of the conclusions of the previous Monte-Carlo runs, as well as the selected scenario of the previous section hold in their totality.

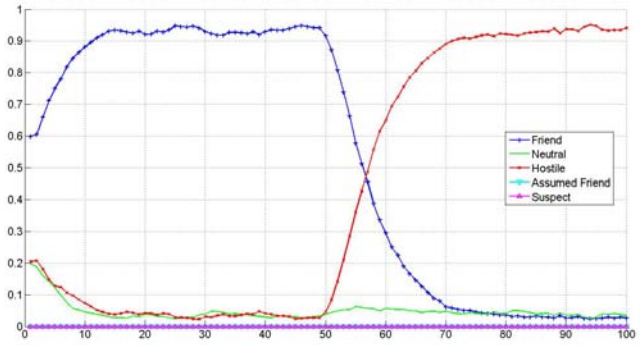


Figure 11: DS result after 100 runs and input filter.

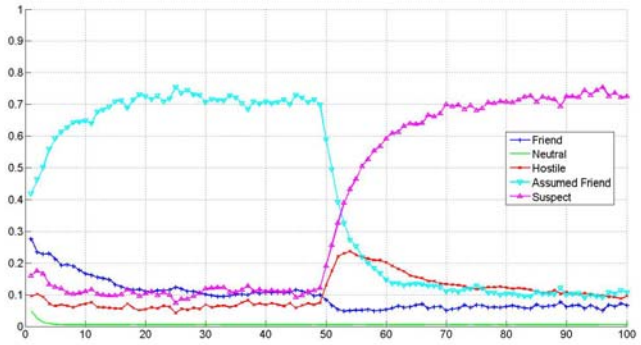


Figure 12: DSm result after 100 runs and input filter.

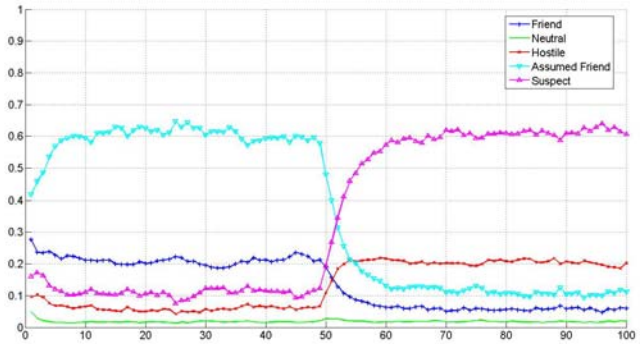


Figure 13: PCR5 result after 100 runs and input filter.

5 Conclusions

Because of the nature of ESM which consists of passive receivers that can identify emitters coming from a small bearing angle, and which, in turn, can be related to platforms that belong to 3 classes: either Friend, Neutral, or Hostile, and to the fact that decision makers would prefer results presented in STANAG 1241 allegiance form, which adds 2 new classes: Assumed Friend, and Suspect, Dezert-Smarandache theory was used instead, but also compared to Dempster-Shafer theory. In Dezert-Smarandache theory an intersection of Friend and Neutral can lead to an Assumed Friend, and an intersection of Hostile and Neutral can lead to a Suspect.

Results were presented showing that the theory can be successfully applied to the problem of associating ESM reports to established tracks, confirming the work published in [15]. Results are also compared to Dempster-Shafer theory which can only reason on the original 3 classes. Thus decision makers are offered STANAG 1241 allegiance results in a timely manner, with quick allegiance change when appropriate, and stability in allegiance declaration otherwise.

In more details, results were presented for a typical scenario and for Monte-Carlo runs with the same conclusions, namely that Dempster-Shafer works well over the original 3 classes, if a minimum to the ignorance is applied. The same can be said for Dezert-Smarandache theory, and to a lesser extent for a popular Proportional Conflict Redistribution rule, but with the added benefit that Dezert-Smarandache theory identifies when miss-associations occur, and to what extent.

Finally, the effects of varying the input parameters for the performance of the ESM were studied, and all of the conclusions remain the same.

References

[1] STANAG 1241, NATO Standard Identity Description Structure for Tactical Use, North Atlantic Treaty Organization, April 2005.

[2] Dempster, A.P., "Upper and lower probabilities induced by a multivalued mapping", Ann. Math. Statist. 38 pp. 325–339 (1967).

[3] Shafer G. *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[4] STANAG 4162, Technical Characteristics of the NATO Identification System (NIS), March 2000.

[5] Henrich, W., Kausch, T., & Opitz, F., "Data Fusion for the new German F124 Frigate Concept and Architecture", *6th International Conference on Information Fusion, FUSION 2003*, Cairns, Queensland, Australia, 8-11

July 2003, CD-ROM ISBN 0-9721844-3-0, and paper proceedings, pp. 1342-1349.

[6] Henrich, W., Kausch, T., & Opitz, F., "Data Fusion for the Fast Attack Craft Squadron 2000: Concept and Architecture", *7th International Conference on Information Fusion, FUSION 2004,* Stockholm, Sweden, 29 June to 1 July 2004, CD-ROM ISBN 91-7170-000-00, and at http://www.fusion2004.foi.se/papers/IF04-0842.pdf .

[7] Valin, P. and Boily, D., "Truncated Dempster-Shafer Optimization and Benchmarking", *Sensor Fusion: Architectures, Algorithms, and Applications IV, SPIE Aerosense 2000*, Orlando, April 24-28 2000, Vol. 4051, pp. 237-246,

[8] Smets Ph., "Data Fusion in the Transferable Belief Model", *3rd International Conference on Information Fusion, Fusion 2000,* Paris, July 10-13, 2000, pp. PS21-PS33.

[9] Multinational Maritime Tactical Instructions and Procedures, MTP 1(D), Vol I, Chapter 6, January 2000.

[10] Smarandache, F., Dezert, J. editors, *Advances and Applications of DSmT for Information Fusion, vol. 1,* American Research Press, 2004.

[11] Smarandache, F., Dezert, J. editors, *Advances and Applications of DSmT for Information Fusion, vol. 2,* American Research Press, 2006.

[12] Simard M.A., Valin P. and Shahbazian E., "Fusion of ESM, Radar, IFF and other Attribute Information for Target Identity Estimation and a Potential Application to the Canadian Patrol Frigate", *AGARD 66th Symposium on Challenge of Future EW System Design,* 18-21 October 1993, Ankara (Turkey), AGARD-CP-546, pp. 14.1-14.18.

[13] Djiknavorian, P., "Fusion d'informations dans un cadre de raisonnement de Dezert-Smarandache appliquée sur des rapports de capteurs ESM sous le STANAG 1241", Mémoire de maîtrise, Université Laval, 2008.

[14] Bieker, T., "Statistical Evaluation of Decision-Level Fusion Methods for Non-Cooperative Target Identification by Radar Signatures", *11th International Conference on Information Fusion, FUSION 2008,* Cologne, Germany, June 30- July 03 2008.

[15] Djiknavorian, P., Grenier, D., and Valin, P., "Analysis of information fusion combining rules under the DSm theory using ESM input", *10th International Conference on Information Fusion, FUSION 2007,* Québec, Canada, 9-12 July 2007.

An Application of *DSmT* in Ontology-Based Fusion Systems

Ksawery Krenc
Adam Kawalec

Originally published as Krenc K., Kawalec A., *An application of DSmT in ontology-based fusion systems*, in Proc. of Fusion 2009, Seattle, WA, USA, 6-9 July 2009, and reprinted with permission.

Abstract – *The aim of this paper is to propose an ontology framework for preselected sensors due to the sensor networks' needs, regarding a specific task, such as the target's threat recognition. The problem will be solved methodologically, taking into account particularly non-deterministic nature of functions assigning the concept and the relation sets into the concept and relation lexicon sets respectively and vice-versa. This may effectively enhance the efficiency of the information fusion performed in sensor networks.*

Keywords: Attribute information fusion, *DSmT*, belief function, ontologies, sensor networks.

1 Introduction

Ontologies of the most applied sensors do not take into account needs of sensor networks [1]. Sensors, in particular the more complex ones, like radars or sonars are intended to be utilized autonomously.

The foundation of the sensor networks (*SN*), comprehended as the networks of cooperative monitoring, is *understanding* information obtained from some elements by another ones. Thus the question of *the common language* is very important. The ontology of sensor network should be unified and structured.

The key problem in this paper is neither a direct application of existing solutions in the field of ontologies for the sensor networks nor a design of a new ontology, ready to implement. The aim is to propose the ontology framework for networks, consisting of preselected sensors, due to the sensory needs, to perform a specific task, such as recognizing the target threat.

The selection of the sensors will be taken in four particular steps, namely:

1. Describing, what particular pieces of information are required to define the target threat;

2. Describing, what particular sensors enable to gain the mentioned pieces of information;

3. Identification of all information possible to acquire by preselected sensors;

4. The specific sensor selection;

2 Sensor type selection

This section focuses on creating the ontology of a sensor network, processing information related to the target threat attribute. Mentioned information may be classified, according to its origin, as:

- Observable – originated directly from sensors or visual sightings;
- Deductable (abductable) – designated by the way of deductive reasoning, based on the other observable attributes, gathered previously;
- Observable and deductable – designated both: on the basis of observation and by the way of deductive reasoning;
- Confirmed – verified by other information center or external sensor network;

The observable attributes may be defined based on information originated from diverse sensors. For the purpose of this paper the scope of sensors (possible to utilize) will be constrained to the set, which in the authors' opinion fully reflects the required information about the target in the real world.

It is a very important assumption that the selection of sensor types is conditioned ontologically. That means neither any particular sensor model nor communication protocol nor any other element of the *SN* organizations will be discussed.

From the observer's point of view (whose main duty is to assess the target threat) it is important to define the following features of the target:

- Key attribute of the target: the threat (based on observations);
- Additional target attributes (as the basis for deduction reasoning about the threat) i.e. the platform, (frigate, corvette, destroyer) and the activity (attack, reconnaissance, search & rescue);
- Auxiliary characteristics: target position;

2.1 Types of sensors

Preselected target features may be registered by various means of observation, namely:

- Position: Radar (all spatial dimensions), sonar, IR sensor (mostly to define target azimuth and elevation);
- Threat: IFF, visual sightings (human), video camera (daylight or noctovision);
- Platform: visual sightings, video camera, thermo-vision camera;
- Activity: visual sightings.

The above statement may be regarded as a pre-selection of sensor set, used in the following considerations of this paper. It is important to notice, that some of the mentioned sensors may acquire information related to more than one attribute. Therefore, a reversed assignment (sensors to attributes) seems to be more adequate.

2.2 Sensor-originated information

Figure 1 presents the preselected target features and their inclusion relations. Additionally, it was pointed out the example sensors, which enable to acquire the mentioned information.

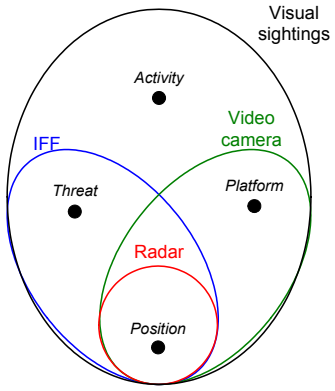


Figure 1 Information scope originated from diverse types of sensors.

It should be noted that although some of these sources allow for obtaining information on more than one attribute, it is possible to identify a hierarchy of relevance of this information. That means that some of the attributes, however, possible to reveal from multiple sources, for some sources perform the primary information while for others the secondary information:

- Radar: position[1];
- IFF: position, threat;
- Video camera: position, platform, threat;
- Visual sightings: position, threat, platform, activity;

For visual sightings, where the human plays the role of the sensor, it is difficult to identify the primary information. Among the above sources the visual recognition is the most reliable way of defining the target activity. Therefore, taking into account the fact that it allows to identify the target threat and platform, the visual recognition may be considered as a specific source of information.

These observations are highly important for future considerations, which will be effectively used in creation of the hierarchy of the concept lexicons as well as in defining the relations among concepts of *SN* ontology.

Some of these sensors perform very complex devices and require the introduction of certain interfaces, allowing the automatic acquisition of useful information (in terms of sensor networks). An example of such a sensor is a video camera. In order to make effective use of an image from the video camera a specific module is necessary to interpret the taken picture, identifying the significant features of the object of interest. In that case, the ontology, the video camera is defined in that very module and it is modifiable as long as there is access to the configuration of that module. This leads to another possible classification of sensors:

- Constant (invariant) ontology sensors, e.g. IFF;
- Variant ontology sensors, e.g. video camera equipped with interpretation module or visual sightings;

Guided by the principle of maximum information growth, in next stages of creating the *SN* ontology the following sources of attribute information will be taken into account: IFF, video camera (*VC*) and visual sightings (*VS*).

3 Defining sets of *SN* ontologies

Referring to a taxonomy of the term of *ontology* [1] the authors would like to notice that the problem of *SN* ontology concerns, in particular, the so-called *method and task ontologies*.

There have been effectively utilized concept lexicons of *Joint C3 Information Exchange Data Model*

[1] Underline means the prime information.

[2], constraining the considerations to three of the *JC3* model attributes:

- threat: *object-item-hostility-status-code*;
- platform: *surface-vessel-type-category-code*;
- activity: *action-task-activity-code*;

While defining the attribute relation functions, the Dezert-Smarandache Theory (*DSmT*) of plausible and paradoxical reasoning has been utilized [3].

3.1 Rules for sensor network ontologies selection

In section 2.2 there was proposed a sensor distinction for variant and invariant ontology sensors. Considering this division is fundamental while creating *SN* ontology, which takes place in four stages:

1. Creating the fundamental concept lexicon for a sensor network, based on invariant concept lexicons of particular sensors;
2. Creating the auxiliary concept lexicon for sensor network, based on variant concept lexicons of particular sensors;
3. Extending the fundamental concept lexicon with the auxiliary lexicon;
4. Defining relations among the concepts in sensor network;

According to the definition of ontology, given in [4], [5], *SN* ontology may be formulated as follows:

$$O = \left\langle L, F, G, \overline{F}, \overline{G}, C, R \right\rangle \tag{1}$$

where:

L – is either concept or relation lexicon; F – lexicon elements to concepts assigning function; G – lexicon elements to relations assigning function; $\overline{F}$ – a function reversed to F, assigning concepts to elements of the concept lexicon; $\overline{G}$ – a function reversed to G, assigning relations to elements of the relation lexicon; C – a set of the whole concepts used in *SN*; R – a set of the whole relations used in *SN*.

According to the lexicons of *JC3* model, the above mentioned concepts and functions will be defined in the following subsections.

3.1.1 Concepts

Concepts are representations of a certain group of objects of the same characteristics, which may be directly identified by selected subset of elements of the concept lexicon [5]. That means, that assigning for example an attribute 'hostile target' to a target uses the concept of the 'hostile target', which is the element of the set (C) of all possible concepts for a given sensor network.

Another question is a representation of the concept 'hostile target' in the language of the particular source. For instance: for IFF device it will be the value of 'FOE', and for a video camera the value, defined in the interpreting module as 'HOSTILE'.

Mathematically, the F assignment is not a bijection in general, moreover: it is not a function. In case multiple sources are utilized, the F is not an injection, whereas if the concept set is 'rich', comparably to the 'poor' lexicon the $\overline{F}$ is not injective. This may occur if the *SN*, prepared for defining fully target threat, is used for deciding whether the target is either friend or hostile. Then, the $\overline{F}$ will interpret concepts of 'training hostile', 'training suspect' and 'assumed friend' as 'friend' assigning the lexical value of 'FRIEND' [6].

In order to illustrate F and $\overline{F}$ assignment it is suggested to consider the following example.

Example 1: Let the set of concepts be defined as follows:

$$C = \{\text{'friend'}, \text{'assumed friend'}, \text{'assumed hostile'}, \text{'hostile'}\} \tag{2}$$

and the concept lexicon is defined as follows:

$$Lc = \{FRIEND, HOSTILE, ASSUMED\} \tag{3}$$

Thus, it is possible to define subsets of the concept lexicon elements in such a way that the F assignment would be a bijection (Figure 2).

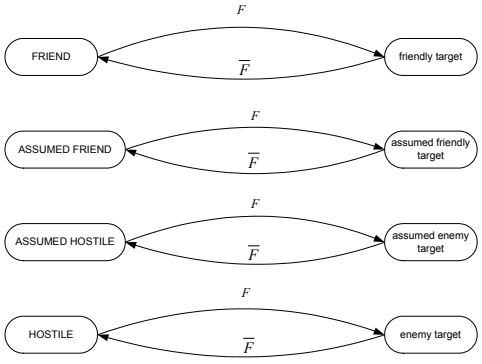


Figure 2 F-assignment as a bijection.

Defining subsets of lexical elements as singletons leads to non-function *F* assignment (Figure 3).

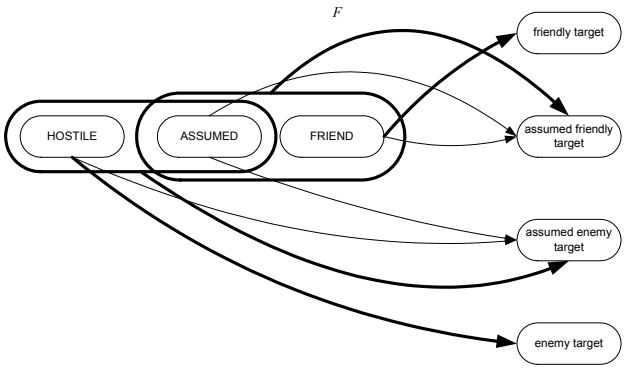


Figure 3 *F* as a non-function assignment.

In case of 'rich' concept lexicon sets it is important to express subsequent target types as conjunctions of their distinctive features.

Example 2:

Table 1 Example definitions of surface platforms

Transporter	AUX ∧ AIR ∧ D ∧ TRAN
Command	AUX ∧ S&MCAL ∧ AIR ∧ C2

where:

AUX – auxiliary vessel;
S&MCAL – equipped with artillery of small and medium caliber;
AIR – against the air targets;
D – performs landing operations;
C2 – command & control;
TRAN – transport of landing forces;

3.1.2 Relations

Relations define the relationships among concepts. Relation may be hierarchical or structural. Moreover, for the purpose of sensor networks, they may be classified as:

- Relations I, among the observable attributes of a diverse type;
- Relations II, among attributes of miscellaneous origin;
- Relations III, among the identical attributes, originated from diverse sources;

Relations among the observable attributes of a diverse type enable a deduction of some attributes values based on observable values of another ones. For instance: the relations between the threat and the platform of the target enable the deduction of target activity. Linking the subsequent observable attributes is performed according to mentioned in previous section distinctive features of the target. This means that for example: defining (based on observations) the target platform is equal to assigning to the target some of distinctive features, which the target, performing the particular activity, has to possess.

Relations among attributes of miscellaneous origin: observable and deductable result in so-called observable-deductable attribute. The effective information fusion from multiple sources is performed according to the rules of combination and conditioning, obtained from *DSmT* [7], [8]. This process is going to be described in details in section 3.2.

Relations among the identical attributes, originated from diverse sources are the type of relations, where the key question is a lexical variety of concepts used by particular sources. For instance: the threat attribute value acquired from IFF may be either FRIEND or FOE, whereas the same attribute obtained from visual sightings may be of {FRIEND, HOSTILE, UNKNOWN, JOKER, FAKER,…}. In such a case a value of FRIEND, gained from IFF, corresponds to the exact value of the visual sightings. The value of FOE is equal to HOSTILE, whereas the relations among values of FRIEND, gained from IFF and FAKER (or JOKER), gained from the visual sightings are not so obvious and they must be defined, according to the definitions of these training types (JOKER, FAKER).

3.2 Proposition of sensor network ontology

This section presents a proposition of an ontology framework for a sensor network, dedicated to monitor the target threat. In the solution there were utilized concepts and concept lexicons of *JC3* model. The authors' intention was to show the way relations of three attributes (threat, platform and activity) should be defined, rather than to present the complete *SN* ontology.

Table 2 presents a bijective assignment of concepts to elements of a concept lexicon. As it was mentioned before, this assignment need not be a bijection, however it is desirable especially if sets of values for attributes of platform and activity are numerous.

Table 2 *SN* ontology: concepts and concept lexicon.

Concepts		**Concept lexicon**	
Threat	An OBJECT-ITEM that is assumed to be a friend because of its characteristics, behavior or origin.	***object-item-hostility-status-code***	ASSUMED FRIEND
	An OBJECT-ITEM that		HOSTILE

	is positively identified as enemy.		
	…according to JC3		… according to JC3
Platform	General designator for aircraft/multi-role aircraft carrier;	***surface-vessel-type-category-code***	AIRCRAFT CARRIER, GENERAL
	Craft 40 meters or less employed to transport sick/wounded and/or medical personnel.		AMBULANCE BOAT
	… according to JC3		… according to JC3
Activity	To fly over an area, monitor and, where necessary, destroy hostile aircraft, as well as protect friendly shipping in the vicinity of the objective area.	***action-task-activity-code***	PATROL, MARITIME
	Emplacement or deployment of one or more mines.		MINE-LAYING
	… according to JC3		… according to JC3

The assignment of relations among attributes to relation lexicons (Table 3) is a surjection. In order to define the relations among attributes *DSmT* combining and conditioning rules have been applied. The preferred rule for conditioning is the rule no. 12. When combining evidence, there is a possibility to use many combination rules, depending the particular relation. However, for simplicity, it is suggested to apply the classic rule of combination (*DSmC*), which has properties of commutativity and associativity.

Table 3 *SN* ontology: relations and relation lexicon.

Relations		**Remarks**	**Relation lexicon**
Rel. I:	*cond(.)*	Based on *DSmT*	*Conditioning*
	$\rightarrow$	According to distinctive features	*Implication*
Rel. II:	*cond(.)*	Based on *DSmT*	*Conditioning*
	$\oplus$	Based on *DSmT*	*Combination*
Rel. III:	*cond(.)*	Based on *DSmT*	*Conditioning*
	$\oplus$	Based on *DSmT* (combination rule need not be identical with one in Relations II)	*Combination*

Below, there have been presented examples of particular types of relations. In case of the relation of type I it is possible to reason about a value of a certain attribute, based on the knowledge about the other ones. However, if the unambiguous deduction of the third attribute is not possible, due to the majority of possible solutions, an application of abductive reasoning (selection of the optimal variant) seems to be justified.

Relations I:

(Threat, Platform) $\rightarrow$ *Activity: (FAKER, FRIGATE TRAINING)* $\rightarrow$ *TRAIN OPERATIONS*;
(Threat, Activity) $\rightarrow$ *Platform: (FAKER, TRAIN OPERATIONS)* $\rightarrow$ *TRAINING CRAFT*;
(Platform, Activity) $\rightarrow$ *Threat: (HOUSEBOAT, PROVIDE CAMPS)* $\rightarrow$ *NEUTRAL*;

Relations II:

$$FAKER = cond(obs(FAKER) \oplus ded(FAKER) \oplus obs(FRIEND))$$;

Relations III:

$$FAKER = cond(obs(FAKER) \oplus VS(FAKER) \oplus IFF(FRIEND))$$;

The abductive reasoning process may be systemized by application of *DSmT*, where the selection of the optimal value takes place after calculating the basic belief assignment.

Example 3:

(Threat, Activity) $\rightarrow$ *Platform: (FRIEND, MINE HUNTING MARITIME)* $\rightarrow$
MINEHUNTER COASTAL (MHC) $\vee$
MINEHUNTER COASTAL WITH DRONE (MHCD) $\vee$ *MINEHUNTER GENERAL (MH)* $\vee$
MINEHUNTER INSHORE (MHI) $\vee$
MINEHUNTER OCEAN (MHO) $\vee$
MINEHUNTER/SWEEPER COASTAL (MHSC) $\vee$
MINEHUNTER/SWEEPER GENERAL (MHS) $\vee$
MINEHUNTER/SWEEPER OCEAN (MHSO) $\vee$
MINEHUNTER/SWEEPER W/DRONE (MHSD)

Applying *DSmT*, for each of possible hypothesis a certain mass of belief is assigned, e.g.:

m(MHC) = 0.2, m(MHCD) = 0.3, m(MH) = 0.1, m(MHI) = 0.1 ,m(MHO) = 0.1, m(MHSC) = 0.05, m(MHS) = 0.05, m(MHSO) = 0.05, m(MHSD) = 0.05

Based on the obtained *basic belief assignment* (*bba)* belief functions, referring to particular hypotheses, may be calculated. In the simplest case, assuming all of the hypotheses are exclusive, the subsequent belief functions will be equal to respective masses, e.g. *Bel(MHC) = m(MHC), Bel(MHCD) = m(MHCD)*, etc.

More complex case, where relationships among hypotheses are taken into account will be considered in the next section.

4 Verification of the usefulness of elaborated ontology sets

The presented framework of the *SN* ontology, for the purpose of the target threat assessment, requires a verification. Particularly, it is important to verify the correctness of reasoning processes and a combination of the reasoning results with observation information.

The proposed solution substantially differs from the existing deterministic ontology-based methods because it introduces explicitly the uncertainty of the relations among target attributes. Therefore this section was meant to focus on the verification of these relation reasoning mechanisms rather than the completeness of the target representation by the sensor network.

4.1 Assumptions

In order to verify the usefulness of the proposed ontology framework, a specially designed demonstrator application for evaluation of the target threat information has been used. This application enables a simulation of acquiring of information from diverse sources, like: radar, video camera and visual sightings.

It is assumed that the visual sighting is also a source of information about a target platform and a target activity. The *bba* values for platform and activity attributes have been assigned arbitrary. During experimentation the observable attributes as well as deductable attributes have been taken into account. Frames of discernment for observation and deduction may differ in general. For the purpose of verification of proposed ontology sets, an example from the section 3.2 is to be considered. Additionally it is assumed:

- Application of the hybrid *DSmT* model:
  - The hypotheses are not exclusive;
  - The hypotheses correspond to the JC3 model terminology;
- In relations of type II and III the hybrid rule of combination (*DSmH)* has been applied;
- The conditioning rule no. 12 has been used for updating evidences;

4.2 Numerical experiments

Figure 4 shows a randomly generated trajectory of the target of which the threat value is at stake. Observations are taken from three sources (visual sightings, radar system - IFF and video camera) synchronously.

The green color means successively acquired observations for each of the sources. The red color means the observations impossible to acquire because the target was outside of the detection zone for a particular source [3].

Taking for example the last sample, the respective *bba* are as Table 4 shows.

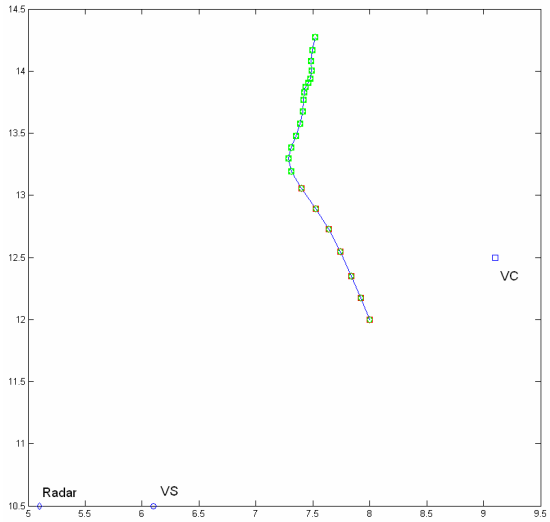


Figure 4 Randomly generated target trajectory and its threat evaluation based on radar, *VS* and *VC* observations.

Table 4 *Bba* gathered from diverse sources: visual sightings, video camera and radar.

Threat	Visual Sightings	Video Camera	Radar/IFF
HOS	0.0024	0.0004	0.0008
UNK	0.0060	0.0012	-
NEU	0.0068	0.0015	-
JOK	0.0109	-	-
FRD	0.2400	0.4368	0.8773
FAK	0.0292	0.0049	0.0119
SUS	0.0032	0.0005	0.0011
AFR	0.0215	0.0046	0.0088
PEN	0.6800	0.5500	0.1000

A relation of type III of combining information from IFF and the visual sightings results in acceptance the target is friendly:

$$Threat_{VS} \oplus Threat_{IFF} \equiv FRIEND \qquad (4)$$

From the visual sightings it is also acquired that the target activity is mine-hunting (*MINE HUNTING MARITIME*). Thus, the relation of type I, between the threat and the activity attribute results in selection of the target platform, related to searching for mines.

$$(FRIEND,\ MINE\ HUNTING\ MARITIME) \rightarrow platform \qquad (5)$$

In the considered case it is assumed the frame of discernment of the *platform* attribute originated from the video camera is defined as follows:

$$\Theta_{VC} = \{MHC, MHI, MHO, MSC, MSO, D\} \quad (6)$$

where:

MHC – MINEHUNTER COASTAL;
MHI – MINEHUNTER INSHORE;
MHO – MINEHUNTER OCEAN;
MSC – SWEEPER COASTAL;
MSO – SWEEPER OCEAN;
D – DRONE;

Additionally, with $\cup$ and $\cap$ operators the secondary hypotheses may be created, which refer to another values of the *platform* attribute (*surface-vessel-type-category code*) of *JC3* model:

MHC $\cup$ *D = MHCD (MINEHUNTER COASTAL WITH DRONE);*

MHI $\cup$ *MHO* $\cup$ *MHC* $\cup$ *D = MH (MINEHUNTER GENERAL);*

MHO $\cap$ *MSO = MHSO (MINEHUNTER/SWEEPER OCEAN);*

(MHC $\cap$ *MSC)* $\cup$ *D = MHSD (MINEHUNTER/SWEEPER W/DRONE);*

(MHO $\cap$ *MSO)* $\cup$ *(MHC* $\cap$ *MSC)* $\cup$ *D = MHS (MINEHUNTER/SWEEPER GENERAL);*

The basic belief assignment for the video camera observation may be defined as follows:

$m_{VC}(MHC) = 0.1$, $m_{VC}(MHCD) = 0.1$,
$m_{VC}(MSC) = 0.2$, $m_{VC}(MHI) = 0.3$,
$m_{VC}(MHO) = 0.2$, $m_{VC}(MSO) = 0.1$,

Due to the implication (5) the above *bba* may be modified according to BCR12 with a following condition:

$$Cond : Truth = MHC \cup MHO \cup MHI \quad (7)$$

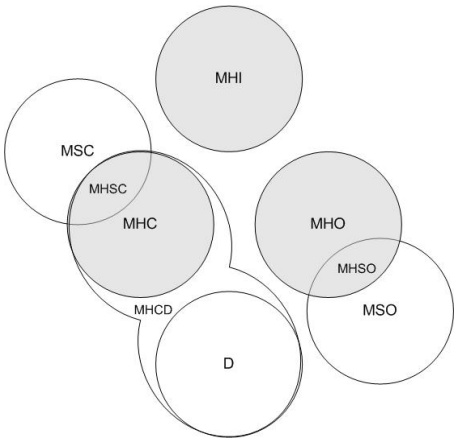


Figure 5 Venn's diagram for the platform attribute. The truth is grey colored.

Thus, the resulting *bba* for the platform attribute is updated, as follows:

$m_R(MHC|Cond)=m_{VC}(MHC)+m_{VC}(MHCD)=0.2$,
$m_R(MHSC|Cond)=m_{VC}(MSC)=0.2$,
$m_R(MHI|Cond)=m_{VC}(MHI)=0.3$,
$m_R(MHO|Cond)=m_{VC}(MHO)=0.2$,
$m_R(MHSO|Cond)=m_{VC}(MSO)=0.1$,

which, after calculating the respective belief and plausibility functions, leads to acceptation of the hypothesis of *MHC (MINEHUNTER COASTAL)* for the platform attribute of the whole sensor network.

It is worth of notice that the belief function for *MHC* before updating is of the least value since:

$$Bel_{VC}(MHC) = m_{VC}(MHC) = 0.1 \quad (8)$$

After updating, due to the fact that $m_{VC}(MHSC)$ supports the belief in *MHC* hypothesis, this hypothesis becomes the most credible since:

$$Bel_R(MHC) = m_R(MHC) + m_R(MHSC) = 0.4 \quad (9)$$

5 Conclusions

The results of the numerical experiments, presented in the previous section, have proven that the application of *DSmT* for the purpose of defining relations among target attributes, gives the possibility of unification of information acquired from sensors as well as obtained based on the deductive reasoning. That influences effectively the whole *SN* ontology, due to the fact the *SN* concept lexicon becomes substantially modified. It does not provide a union of lexicons for each sensor, which would be expectable in the deterministic case. The *SN* concept lexicon becomes extended with intersections and unions of the hypotheses created upon the lexicons of particular sensors.

During the experiments it has been utilized the *JC3* model's lexicon of *surface-vessel-type-category-code* attribute. It is important to notice, that despite its large volume, the lexicon is not structured. Thus, an emerging conclusion occurs, that setting *JC3* lexicons in a hierarchy would bring tangible benefits due to the fact that the hierarchy enables creating the hypotheses using $\cup$ and $\cap$ operators more effectively, and this in turn increases the precision of the reasoning processes based on information acquired from sensors.

Acknowledgements

The research work described in this document is a part of extensive works devoted to sensor networks in NEC environment. It was financed with science means from 2007 to 2010 as an ordered research project.

References

[1] K. Krenc: *An introductory analysis of the usefulness of sensor network organizations,* MCC Conference, Cracow, ISBN 83-920120-5-4, 2008.

[2] The Joint C3 Information Exchange Data Model, Edition 3.1b, 2007

[3] K. Krenc, A. Kawalec: *An evaluation of the attribute information for the purpose of DSmT fusion in C&C systems*, Fusion2008, Cologne, ISBN 978-3-00-024883-2, 2008.

[4] http://www.aifb.uni-karlsruhe.de/WBS/meh/publications/ehrig03ontology.pdf

[5] M. Chmielewski, R. Kasprzyk: *Usage and characteristics of ontology models in Network Enabled Capability operations*, MCC Conference, Cracow, ISBN 83-920120-5-4, 2008.

[6] NATO Standardization Agency, *Tactical Data Exchange – Link 16*, STANAG No. 5516, Ed. 3.

[7] Florentin Smarandache, Jean Dezert, *Advances and Applications of DSmT for Information Fusion*, Vol 1, American Research Press Rehoboth, 2004.

[8] Florentin Smarandache, Jean Dezert, *Advances and Applications of DSmT for Information Fusion*, Vol 2, American Research Press Rehoboth, 2006.

GMTI and IMINT Data Fusion for Multiple Target Tracking and Classification

Benjamin Pannetier
Jean Dezert

Originally published as Pannetier B., Dezert J., *GMTI and IMINT data fusion for multiple target tracking and classification*, in Proc. of Fusion 2009, Seattle, WA, USA, 6-9 July 2009, and reprinted with permission.

Abstract − *In this paper, we propose a new approach to track multiple ground target with GMTI (Ground Moving Target Indicator) and IMINT (IMagery INtelligence) reports. This tracking algorithm takes into account road network information and is adapted to the out of sequence measurement problem. The scope of the paper is to fuse the attribute type information given by heterogeneous sensors with DSmT (Dezert Smarandache Theory) and to introduce the type results in the tracking process. We show the ground target tracking improvement obtained due to better targets discrimination and an efficient conflicting information management on a realistic scenario.*

Keywords: Multiple target tracking, heterogeneous data fusion, DSmT.

1 Introduction

Data fusion for ground battlefield surveillance is more and more strategic in order to create the situational assessment or improve the precision of fire control system. The challenge of data fusion for the theatre surveillance operation is to know where the targets are, how they evolve (manoeuvres, group formations,...) and what are their identities.

For the first two questions, we develop new ground target tracking algorithms adapted to GMTI (Ground Moving Target Indicator) sensors. In fact, GMTI sensors are able to cover a large surveillance area during few hours or more if several sensors exists. However, ground target tracking algorithms are used in a complex environment due to the high traffic density and the false alarms that generate a significant data quantity, the terrain topography which can provocate occlusion areas for the sensor and the high maneuvrability of the ground targets which yields to the data association problem. Several references exist for the MGT (Multiple Ground Tracking) with GMTI sensors [1, 2] whose fuse contextual informations with MTI reports. The main results are the improvement of the track precision and track continuity. Our algorithm [6] is built with several reflexions inspired of this literature. Based on road segment positions, dynamic motion models under road constraint are built and an optimized projection of the estimated target states is proposed to keep the track on the road. A VS-IMM (Variable Structure Interacting Multiple Models) filter is created with a set of constrained models to deal with the target maneuvers on the road. The set of models used in the variable structure is adjusted sequentially according to target positions and to the road network topology.

Now, we extended the MGT with several sensors. In this paper, we first consider the centralized fusion between GMTI and IMINT (IMagery INTelligence) sensors reports. The first problem of the data fusion with several sensors is the data registration in order to work in the same geographic and time referentials. This point is not presented in this paper. However, in a multisensor system, measurements can arrive out of sequence. Following Bar-Shalom and Chen's works [3], the VS-IMMC (VS-IMM Constrained) algorithm is adapted to the OOSM (Out Of Sequence Measurement) problem, in order to avoid the reprocessing of entire sequence of measurements. The VS-IMMC is also extended in a multiple target context and integrated in a SB-MHT (Structured Branching - Multiple Hypotheses Tracking). Despite of the resulting track continuity improvement for the VS-IMMC SB-MHT algorithm, unavoidable association ambiguities arise in a multi-target context when several targets move in close formation (crossing and passing). The associations between all constrained predicted states are compromised if we use only the observed locations as measurements. The weakness of this algorithm is due to the lack of good target state discrimination.

One way to enhance data associations is to use the reports classification attribute. In our previous work [5], the classification information of the MTI segments has been introduced in the target tracking process. The idea was to maintain aside each target track a set of ID

hypotheses. Their committed belief are revised in real time with the classifier decision through a very recent and efficient fusion rule called proportional conflict redistribution (PCR). In this paper, in addition to the measurement location fusion, a study is carried out to fuse MTI classification type with image classification type associated to each report. The attribute type of the image sensors belongs to a different and better classification than the MTI sensors. The counterpart is the short coverage of image sensors that brings about a low data quantity. In section 2, the motion and measurement models are presented with a new ontologic model in order to place the different classification frames in the same frame of discernment. After the VS-IMMC description given in section 3, the PCR fusion rule originally developed in DSmT (Dezert-Smarandache Theory) framework is presented in section 4 to fuse the target type information available and to include the resulting fused target ID into the tracking process. The last part of this paper is devoted to simulation results for a multiple target tracking scenario within a real environment.

2 Motion & observation models

2.1 GIS description

The GIS (Geographical Information System) used in this work contains both the segmented road network and the DTED (Digital Terrain Elevation Data). Each road segment expressed in WGS84 is converted in a Topographic Coordinate Frame (denoted *TCF*). The *TCF* is defined according to the origin O in such a way that the axes X, Y and Z are respectively oriented towards the local East, North and Up directions. The target tracking process is carried out in the *TCF*.

2.2 Constrained motion model

The target state at the current time t_k is defined in the local horizontal plane of the *TCF*:

$$\mathbf{x}(k) = [x(k)\ \dot{x}(k)\ y(k)\ \dot{y}(k)]' \tag{1}$$

where $(x(k), y(k))$ and $(\dot{x}(k), \dot{y}(k))$ define respectively the target location and velocity in the local horizontal plane. The dynamics of the target evolving on the road are modelized by a first-order differential system. The target state on the road segment s is defined by $\mathbf{x}_s(k)$ where the target position $(x_s(k), y_s(k))$ belongs to the road segment s and the corresponding heading $(\dot{x}_s(k), \dot{y}_s(k))$ is in its direction.

The event that the target is on road segment s is noted $e_s(k) = \{\mathbf{x}(k) \in s\}$. Given the event $e_s(k)$ and according to a motion model M_i, the estimation of the target state can be improved by considering the road segment s. It follows:

$$\mathbf{x}_s(k) = \mathbf{F}_{s,i}(\Delta(k)) \cdot \mathbf{x}_s(k-1) + \mathbf{\Gamma}(\Delta(k)) \cdot \mathbf{v}_{s,i}(k) \tag{2}$$

where $\Delta(k)$ is the sampling time, $\mathbf{F}_{s,i}$ is the state transition matrix associated to the road segment s and adapted to a motion model M_i, $\mathbf{v}_{s,i}(k)$ is a white Gaussian random vector with covariance matrix $\mathbf{Q}_{s,i}(k)$ chosen in such a way that the standard deviation along the road segment is higher than the standard deviation in the orthogonal direction. It is defined by:

$$\mathbf{Q}_{s,i}(k) = \mathbf{R}_{\theta_s} \cdot \begin{pmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_n^2 \end{pmatrix} \cdot \mathbf{R}_{\theta_s}^{'} \tag{3}$$

where $\mathbf{R}_{\theta_s}$ is the rotation matrix associated with the direction θ_s defined in the plane (O, X, Y) of the road segment s. The matrix $\mathbf{\Gamma}(\Delta_k)$ is defined in [8].

To improve the modeling for targets moving on a road network, we proposed in [5] to adapt the level of the dynamic model's noise based on the length of the road segment s. The idea is to increase the standard deviation σ_n defined in (3) to take into account the error on the road segment location. After the state estimation obtained by a Kalman filter, the estimated state is then projected according to the road constraint $e_s(k)$. This process is detailed in [6].

2.3 GMTI measurement model

According to the NATO GMTI format [7], the MTI reports received at the fusion station are expressed in the WGS84 coordinates system. The MTI reports must be converted in the *TCF*. A MTI measurement z at the current time t_k is given in the *TCF* by:

$$\mathbf{z}(k) = [x(k)\ y(k)\ \dot{\rho}(k)]' \tag{4}$$

where $(x(k), y(k))$ is the location of the MTI report in the local frame (O, X, Y) and $\dot{\rho}(k)$ is the associated range measurement expressed by:

$$\dot{\rho}(k) = \frac{(x(k) - x_c(k)) \cdot \dot{x}(k) + (y(k) - y_c(k)) \cdot \dot{y}(k)}{\sqrt{(x(k) - x_c(k))^2 + (y(k) - y_c(k))^2}} \tag{5}$$

where $(x_c(k), y_c(k))$ is the sensor location at the current time in the TCF. Because the range radial velocity is correlated to the MTI location components, the use of an extended Kalman filter (EKF) is not adapted. In the literature, several techniques exist to uncorrelate the range radial velocity from the location components. We prefer to use the AEKF (Alternative Extended Kalman Filter) proposed by Bizup and Brown in [9], because the implementation is easier by using the alternative lienarization than another algorithms to decorrelate the components. Moreover, AEKF working in the sensor referential/frame remains invariant by translation. The AEKF measurement equation is given by:

$$\mathbf{z}_{MTI}(k) = \mathbf{H}_{MTI}(k) \cdot \mathbf{x}(k) + \mathbf{w}_{MTI}(k) \tag{6}$$

where $\mathbf{w}_{MTI}(k)$ is a zero-mean white Gaussian noise vector with a covariance $R_{MTI}(k)$ (given in [5]) and

$\mathbf{H}_{MTI}(k)$ is defined by:

$$\mathbf{H}_{MTI}(k) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \frac{\partial \dot{\rho}(k)}{\partial \dot{x}} & 0 & \frac{\partial \dot{\rho}(k)}{\partial \dot{y}} \end{pmatrix} \quad (7)$$

Each MTI report is characterized both with the location and velocity information and also with the attribute information and its probability that it is correct. We denote C_{MTI} the frame of discernment on target ID based on MTI data. C_{MTI} is assumed to be constant over the time and consists in a finite set of exhaustive and exclusive elements representing the possible states of the target classification. In this paper, we consider only 3 elements in C_{MTI} defined as:

$$C_{MTI} = \left\{ \begin{array}{l} c_1 \triangleq \textit{Tracked vehicle} \\ c_2 \triangleq \textit{Wheeled vehicle} \\ c_3 \triangleq \textit{Rotary wing aircraft} \end{array} \right\} \quad (8)$$

We consider also the probabilities $P\{c(k)\}$ ($\forall c(k) \in C_{MTI}$) as input parameters of our tracking systems characterizing the global performances of the classifier. The vector of probabilities $[P(c_1)\ P(c_2)\ P(c_3)]$ represents the diagonal of the confusion matrix of the classification algorithm assumed to be used. Let $\mathbf{z}^{\star}_{MTI}(k)$ the extended MTI measurements including both kinematic part and attribute part expressed by te herein formula:

$$\mathbf{z}^{\star}_{MTI}(k) \triangleq \{\mathbf{z}_{MTI}(k), c(k), P\{c(k)\}\} \quad (9)$$

2.4 IMINT motion model

For the imagery intelligence (IMINT), we consider two sensor types : a video EO/IR sensor carried by a Unanimed Aerial Vehicule (UAV) and a EO sensor fixed on a Unattended Ground Sensor (UGS).

We assume that the video information given by both sensor types are processed by their own ground stations and that the system provides the video reports of target detections with their classification attributes. Moreover, a human operator selects targets on a movie frame and is able to choose its attribute with a HMI (Human Machine Interface). In addition, the operator is able with the UAV to select *several targets* on a frame. On the contrary, the operator selects *only one target* with the frames given by the UGS. There is *no false alarm* and a target cannot be detected by the operator (due to terrain mask for example). The video report on the movie frame is converted in the *TCF*. The measurement equation is given by:

$$\mathbf{z}_{video}(k) = \mathbf{H}_{video}(k) \cdot \mathbf{x}(k) + \mathbf{w}_{video}(k) \quad (10)$$

where $\mathbf{H}_{video}$ is the observation matrix of the video sensor

$$\mathbf{H}_{video} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (11)$$

The white noise Gaussian process $\mathbf{w}_{video}(k)$ is centered and has a known covariance $\mathbf{R}_{video}(k)$ given by the ground station.

Each video report is associated to the attribute information $c(k)$ with its probability $P\{c(k)\}$ that it is correct. We denote C_{video} the frame of discernment for an EO/IR source. As C_{MTI}, C_{video} is assumed to be constant over the time and consists in a finite set of exhaustive and exclusive elements representing the possible states of the target classification. In this paper, we consider only eight elements in C_{video} as follows:

$$C_{video} = \left\{ \begin{array}{c} \textit{civilian car} \\ \textit{military armoured car} \\ \textit{wheeled armoured vehicule} \\ \textit{civilian bus} \\ \textit{military bus} \\ \textit{civilian truck} \\ \textit{military armoured truck} \\ \textit{copter} \end{array} \right\} \quad (12)$$

Let $\mathbf{z}^{\star}_{video}(k)$ be the extended video measurements including both kinematic part and attribute part expressed by the following formula ($\forall c(k) \in C_{video}$):

$$\mathbf{z}^{\star}_{video}(k) \triangleq \{\mathbf{z}_{video}(k), c(k), P\{c(k)\}\} \quad (13)$$

For notation convenience, the measurements sequence $\mathbf{Z}^{k,l}$ represents a possible set of measurements generated by the target up to time k (*i.e.*, there exists a subsequence n and a measurement i such that $\mathbf{Z}^{k,l} = \left\{\mathbf{Z}^{k-1,n}, ..., \mathbf{z}^{\star}_j(k)\right\}$) associated with the track $T^{k,l}$. At the current time k, the track $T^{k,l}$ is represented by a sequence of the state estimates. $\mathbf{z}^{\star}_j(k)$ is the j^{th} measurement available at time k among $m(k)$ validated measurements around the target measurement prediction.

3 Tracking with road constraints

3.1 VS IMM with a road network

The IMM is an algorithm for combining state estimates arising from multiple filter models to get a better global state estimate when the target is under maneuvers. In section 2.2, a constrained motion model i to a road segment s, noted $M^i_s(k)$, was defined. Here we extend the segment constraint to the different dynamic models (among a set of $r+1$ motion models) that a target can follow. The model indexed by $r = 0$ is the stop model. It is evident that when the target moves from one segment to the next, the set of dynamic models changes. In a conventionnal IMM estimator [1], the likelihood function of a model $i = 0, 1, \ldots, r$ is given, for a track $T^{k,l}$, associated with the j-th measurement, $j \in \{0, 1, \ldots, m(k)\}$ by:

$$\Lambda^l_i(k) = p\{\mathbf{z}_j(k) | M^i_s(k), \mathbf{Z}^{k-1,n}\} \quad (14)$$

where $\mathbf{Z}^{k-1,n}$ is the subsequence of measurements associated with the track $T^{k,l}$.

Using the IMM estimator with a stop motion model, the likelihood function of the moving target mode for $i = 1, \ldots, r$ and for $j \in \{0, 1, \ldots, m(k)\}$ is given by:

$$\begin{aligned}\Lambda_i^l(k) = \quad & P_D \cdot p\{\mathbf{z}_j(k)|M_s^i(k), \mathbf{Z}^{k-1,n}\} \cdot (1 - \delta_{m_j,0}) \\ & +(1 - P_D) \cdot \delta_{m_j,0} \end{aligned} \tag{15}$$

while the likelihood of the stopped target mode (*i.e.* $r = 0$) is:

$$\Lambda_0^l(k) = p\{\mathbf{z}_j(k)|M_0^i(k), \mathbf{Z}^{k-1,n}\} = \delta_{m_j,0} \tag{16}$$

where P_D is the sensor detection probability, $\delta_{m_j,0}$ is the Kronecker function defined by $\delta_{m_j,0} = 1$ if $m_j = 0$ and $\delta_{m_j,0} = 0$ whenever $m_j \neq 0$.

The combined/global likelihood function $\Lambda^l(k)$ of a track including a stop model is then given by:

$$\Lambda^l(k) = \sum_{i=0}^{r} \Lambda_i(k) \cdot \mu_i(k|k-1) \tag{17}$$

where $\mu_i(k|k-1)$ is the predicted model probabilities [8].

The steps of the IMM under road segment s constraint are the same as for the classical IMM as described in [8].

In real application, the predicted state could also appear onto another road segment, because of a road turn for example, and we need to introduce new constrained motion models. In such case, we activate the most probable road segments sets depending on the local predicted state $\hat{x}_{i,s}^l(k|k-1)$ location of the track $T^{k,l}$[5, 1]. We consider $r+1$ oriented graphs which depend on the road network topology. For each graph i, $i = 0, 1, \ldots, r$, each node is a constrained motion model M_s^i. The nodes are connected to each other according to the road network configuration and one has a finite set of $r+1$ motion models constrained to a road section. The selection of the most probable motion model set, to estimate the road section on which the target is moving on, is based on a sequential probability ratio test (SPRT).

3.2 OOSM algorithm

The data fusion that operates in a centralized architecture suffers of delayed measurement due to communication data links, time algorithms execution, data quantity,... In order to avoid reordering and reprocessing an entire sequence of measurements for real-time application, the delayed measurements are processed as out-of-sequence measurements (OOSM). The algorithm used in this work is described in [3]. In addition, according to the road network constraint, the state retrodiction step is done on the road.

3.3 Multiple target tracking

For the MGT problem, we use the SB-MHT (Structured Branching Multiple Hypotheses Tracking) presented in [10]. When the new measurements set $\mathbf{Z}(k)$ is received, a standard gating procedure is applied in order to validate MTI reports to track pairings. The existing tracks are updated with VS-IMMC and the extrapolated and confirmed tracks are formed. More details can be found in chapter 16 of [10]. In order to palliate the association problem, we need a probabilistic expression for the evaluation of the track formation hypotheses that includes all aspects of the data association problem. It is convenient to use the log-likelihood ratio (LLR) or a track score of a track $T^{k,l}$ which can be expressed at current time k in the following recursive form:

$$L^l(k) = L^s(k-1) + \Delta L^l(k) \tag{18}$$

with

$$\Delta L^l(k) = \log\left(\frac{\Lambda^l(k)}{\lambda_{fa}}\right) \tag{19}$$

and

$$L(0) = \log\left(\frac{\lambda_{fa}}{\lambda_{fa} + \lambda_{nt}}\right) \tag{20}$$

where λ_{fa} and λ_{nt} are respectively the false alarm rate and the new target rate per unit of surveillance volume and $\Lambda^l(k)$ is the likelihood given in (17).

4 Target type tracking

In [4], Blasch and Kahler fused identification attribute given by EO/IR sensors with position measurement. The fusion was used in the validation gate process to select only the measurement according to the usual kinematic criterion and the belief on the identification attribute. Our approach is different since one uses the belief on the identification attribute to revise the LLR with the *posterior* pignistic probability on the target type. We recall briefly the Target Type Tracking (TTT) principle and explain how to improve VS-IMMC SB-MHT with target ID information. TTT is based on the sequential combination (fusion) of the predicted belief of the type of the track with the current "belief measurement" obtained from the target classifier decision. Results depends on the quality of the classifier characterized by its confusion matrix (assumed to be known at least partially as specified by STANAG). The adopted combination rule is the so-called Proportional Conflict Redistribution rule no 5 (PCR5) developed in the DSmT (Dezert Smarandache Theory) framework since it deals efficiently with (potentially high) conflicting information. A detailed presentation with examples can be found in [12, 11]. This choice is motivated in this typical application because in dense traffic scenarios, the VS-IMMC SB-MHT only based on kinematic information can be deficient during maneuvers and crossroads. Let's recall first what the PCR5 fusion rule

is and then briefly the principle of the (single-sensor based) Target Type Tracker.

4.1 PCR5 combination rule

Let $C_{Tot} = \{\theta_1, \ldots, \theta_n\}$ be a discrete finite set of n exhaustive elements and two distinct bodies of evidence providing basic belief assignments (bba's) $m_1(.)$ and $m_2(.)$ defined on the power-set[1] of C_{Tot}. The idea behind the Proportional Conflict Redistribution (PCR) rules [11] is to transfer (total or partial) conflicting masses of belief to non-empty sets involved in the conflicts proportionally with respect to the masses assigned to them by sources. The way the conflicting mass is redistributed yields actually several versions of PCR rules, but PCR5 (i.e. PCR rule # 5) does the most exact redistribution of conflicting mass to non-empty sets following the logic of the conjunctive rule and is well adapted for a sequential fusion. It does a better redistribution of the conflicting mass than other rules since it goes backwards on the tracks of the conjunctive rule and redistributes the conflicting mass only to the sets involved in the conflict and proportionally to their masses put in the conflict. The PCR5 formula for $s \geq 2$ sources is given in [11]. For the combination of only two sources (useful for sequential fusion in our application) when working with Shafer's model, it is given by $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^{C_{Tot}} \setminus \{\emptyset\}$

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{Y \in 2^{C_{Tot}} \setminus \{X\} \\ X \cap Y = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)}\right] \quad (21)$$

where $m_{12}(X)$ corresponds to the conjunctive consensus on X between the two sources (*i.e.* our a prior bba on target ID available at time $k-1$ and our current observed bba on target ID at time k) and where all denominators are different from zero. If a denominator is zero, that fraction is discarded.

4.2 Principle of the target type tracker

To estimate the true target type $type(k)$ at time k from the sequence of declarations $c(1)$, $c(2)$, $\ldots c(k)$ done by the unreliable classifier[2] up to time k. To build an estimator $\hat{type}(k)$ of $type(k)$, we use the general principle of the Target Type Tracker (TTT) developed in [12] which consists in the following steps:

- a) Initialization step (*i.e.* $k = 0$). Select the target type frame $C_{Tot} = \{\theta_1, \ldots, \theta_n\}$ and set the prior bba $m^-(.)$ as vacuous belief assignment, i.e $m^-(\theta_1 \cup \ldots \cup \theta_n) = 1$ since one has no information about the first observed target type.

- b) Generation of the current bba $m_{obs}(.)$ from the current classifier declaration $c(k)$ based on attribute measurement. At this step, one takes $m_{obs}(c(k)) = P\{c(k)\} = C_{c(k)c(k)}$ and all the unassigned mass $1 - m_{obs}(c(k))$ is then committed to total ignorance $\theta_1 \cup \ldots \cup \theta_n$. $C_{c(k)c(k)}$ is the element of the known confusion matrix $\mathbf{C}$ of the classifier indexed by $c(k)c(k)$.

- c) Combination of current bba $m_{obs}(.)$ with prior bba $m^-(.)$ to get the estimation of the current bba $m(.)$. Symbolically we write the generic fusion operator as $\oplus$, so that $m(.) = [m_{obs} \oplus m^-](.) = [m^- \oplus m_{obs}](.)$. The combination $\oplus$ is done according to the PCR5 rule (*i.e.* $m(.) = m_{PCR5}(.)$).

- d) Estimation of True Target Type is obtained from $m(.)$ by taking the singleton of Θ, *i.e.* a Target Type, having the maximum of belief (or eventually the maximum Pignistic Probability).

$$\widehat{type}(k) = \underset{A \in C_{Tot}}{argmax}(BetP\{A\}) \quad (22)$$

The Pignistic Probability is used to estimate the probability to obtain the type $\theta_i \in C_{Tot}$ given the previous target type estimate $\widehat{type}(k-1)$.

$$BetP\{\theta_i\} = P\{\widehat{type}(k) = \theta_i | \widehat{type}(k-1)\} \quad (23)$$

- e) set $m^-(.) = m(.)$; do $k = k+1$ and go back to step b).

Naturally, in order to revise the LLR in our GMTI-MTT systems for taking into account the estimation of belief of target ID coming from the Target Type Trackers, we transform the resulting bba $m(.) = [m^- \oplus m_{obs}](.)$ available at each time k into a probability measure. In this work, we use the classical pignistic transformation defined by [13]:

$$BetP\{A\} = \sum_{X \in 2^{C_{Tot}}} \frac{|X \cap A|}{|X|} m(X) \quad (24)$$

4.3 Working with multiple sensors

Since in our application, we work with different sensors (i.e. MTI and Video EO/IR sensors), one has to deal with the discernment frames C_{MTI} and C_{video} defined in section 2. Therefore we need to adapt the (single-sensor based) TTT to the multi-sensor case. We first adapt the frame C_{MTI} to C_{video} and then, we extend the principle of TTT to combine multiple bba's (typically here $m_{obs}^{MTI}(.)$ and $m_{obs}^{Video}(.)$) with prior target ID bba $m^-(.)$ to get finally the updated global bba $m(.)$ at each time k. The proposed approach can

[1] In our GMTI-MTT applications, we will assume Shafer's model for the frame C_{Tot} of target ID which means that elements of C_{Tot} are assumed truly exclusive.

[2] Here we consider only one source of information/classifier, say based either on the EO/IR sensor, or on a video sensor by example. The multi-source case is discussed in section 4.3.

be theroretically extended to any number of sensors. When no information is available from a given sensor, we take as related bba the vacuous mass of belief which represents the total ignorant source because this doesn't change the result of the fusion rule [11] (which is a good property to satisfy). For mapping C_{MTI} to C_{video}, we use a (human refinement) process such that each element of C_{MTI} can be associated at least to one element of C_{video}. In this work, the delay on the the type information provided by the video sensor is not taking into account to update the global bba $m(.)$. All type information (delayed or not provided by MTI and video sensors) are considered as bba $m_{obs}(.)$ available for the current update. The explicit introduction of delay of the out of sequence video information is under investigations.

4.4 Data attributes in the VS IMMC

To improve the target tracking process, the introduction of the target type probability is done in the likelihood calculation. For this, we consider the measurement $\mathbf{z}_j^*(k)(\forall j \in \{1,\ldots,m_k\})$ described in (9) and (13). With the assumption that the kinematic and classification observations are independant, it is easy to prove that the new combined likelihood Λ_N^l associated with a track $T^{k,l}$ is the product of the kinematic likelihood (17) with the classification probability in the manner that:

$$\Lambda_N^l(k) = \Lambda^l(k) \cdot P\{\widehat{type}(k)|\widehat{type}(k-1)\} \qquad (25)$$

where the the probability $P\left\{\widehat{type}(k)|\widehat{type}(k-1)\right\}$ is chosen as the pignistic probability value on the declared target type $\widehat{type}(k)$ given $\widehat{type}(k-1)$ derived from the updated mass of belief $m(.)$ according to our target type tracker.

5 Simulations and results

5.1 Scenario description

To evaluate the performances of the VS-IMMC SB-MHT with the attribute type information, we consider 10 maneuvering (acceleration, deceleration, stop) targets on a real road network. The 10 target types are given by (12). The target 1 is passing the military vehicules 2, 3, 4 and 7. Targets 2, 3, 4 and 7 start from the same starting point.The target 2 is passing the vehicules 3 and 7 in the manner that it places in front of the convoy. The targets 5, 6, 9 and 10 are civilian vehicles and are crossing the targets 1, 2, 3 and 7 at several junctions. The goal of this simulation is to reduce the association complexity by taking into account the road network topology and the attribute types given by heterogeneous sensors. In this scenario, we consider one GMTI sensor located at $(-50km, -60km)$ at $4000m$ in elevation and one UAV located at $(-100m, -100m)$ at $1200m$ in elevation and 5 UGS distributed on the ground. The GMTI sensor tracks the 10 targets at every 10 seconds with $20m$, $0.0008rad$ and $1m \cdot s^{-1}$ range, cross-range and range-rate measurements standard deviation respectively. The detection probability P_D is equal to 0.9 and the MDV (Minimal Detectable Velocity) fixed at $1m \cdot s^{-1}$. The false alarms density is fixed ($\lambda_{fa} = 10^{-8}$). The confusion matrix described in part 4.2 is given by:

$$\mathbf{C}_{MTI} = diag(\begin{bmatrix} 0.8 & 0.7 & 0.9 \end{bmatrix}) \qquad (26)$$

This confusion matrix is only used to simulate the target type probability of the GMTI sensor. The data obtained by UAV are given at 10 seconds with $10m$ standard deviation in X an Y direction from the *TCF*. The time delay of the video data is constant and equal to 11 seconds. The detection probability P_D is equal to 0.9. The human operator only selects for each video report a type defined by (12). In our simulations, the target type probability depends on the sensor resolution. For this, we consider the volume V_{video} of the sensor area surveillance on the ground. The diagonal terms of the confusion matrix $\mathbf{C}_{video}$ are equal to $P\{c(k)\}$ where $P\{c(k)\}$ is defined by:

$$P\{c(k)\} = \begin{cases} 0.90 \, if \, V_{video} \leq 10^6 m^2 \\ 0.75 \, if \, 10^6 m^2 < V_{video} \leq 10^8 m^2 \\ 0.50 \, if \, V_{video} > 10^8 m^2 \end{cases} \qquad (27)$$

For the UGS, the target detection is done if only the target is located under the minimal range detection (MRD). The MRD is fixed for the 5 UGS at 1000 m and each sensor gives delayed measurement every seconds. The time delay is also equal to 11 seconds. The UGS specificity is to give only one target detection during 4 seconds in order to detect another target. We recall that there is no false alarms for this sensor. Based on [4], the target type probability depends on α (*i.e.* the target orientation towards the UGS). The more the target orientation is orthogonal to the sensor line of sight, the more the target type probability increases. The diagonal terms of the confusion matrix $\mathbf{C}_{UGS}$ are equal to $P\{c(k)\}$ where $P\{c(k)\}$ is defined by:

$$P\{c(k)\} = \begin{cases} 0.90 & if \, \frac{5\pi}{6} \leq \alpha \leq \frac{\pi}{6} \\ 0.50 & otherwise \end{cases} \qquad (28)$$

For each detected target, a uniform random number $u \sim U([0,1])$ is drawn. If u is greater than the true target type probability of the confusion matrix, a wrong target type is declared for the ID report and used with its associated target type probability. The targets are scanned at different times by the sensors (figure 1).

5.2 Filter parameters

We consider three motion models (*i.e.* $i \in \{0,1,2\}$) which are respectively a stop model M_0 when the target

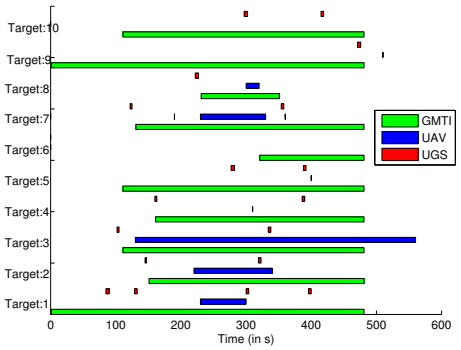


Figure 1: Target's sensor illumination.

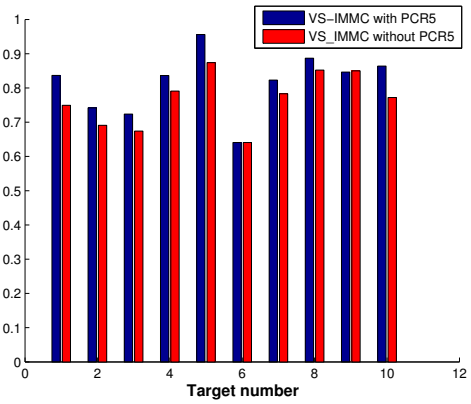


Figure 2: Track length ratio.

is assumed to have a zero velocity, a constant velocity model M_1 with a low uncertainty, and a constant velocity model M_2 with a high uncertainty (modeled by a strong noise). The parameters of the IMM are the following: for the motion model M_1, the standard deviation along and orthogonal to the road segment are equals to $0.05\, m{\cdot}s^{-2}$), the constrained constant velocity model M_2 has a high standard deviation to adapt the dynamics to the target manoeuvre (the standard deviation along and orthogonal to the road segment are respectively equal to $0.8\ m \cdot s^{-2}$ and $0.4\ m \cdot s^{-2}$) and the stop motion model M_0 has a standard deviation equals to zero. These constrained motion models are however adapted to follow the road network topology. The transition matrix and the SB-MHT parameters are those taken in [5].

5.3 Results

For each confirmed track given by the VS-IMMC SB-MHT, a test is used to associate a track to the most probable target. The target tracking goal is to track as long as possible the target with one track. To evaluate the track maintenance, we use the track length ratio criterion, the averaged root mean square error (noted ARMSE) for each target and the track purity and the type purity (only for the tracks obtained with PCR5) [5]. These measures of performances are averaged on 50 Monte-Carlo runs.

On figure 2, one sees that the track length ratio becomes better with the PCR5 than without as expected for the target 6. When the targets 1 and 2 are passing the targets 3, 4 and 7, an association ambiguity arises to associate the tracks with the correct measurements. This is due to the close formation between targets with the GMTI sensor resolution and the road network configureation with junctions. Sometimes tracks are lost with the VS IMMC SB-MHT without the PCR5. Then new tracks for each targets are built. That is why, the track purity of the VS IMMC SB-MHT without PCR5 (Table 1) is smaller than the the track purity with PCR5 (Table 2). So, the track precision, given by the ARMSE criterion, is better with the PCR5. For the target 6 results, this target is only scanned by the GMTI sensor and its associated performances are equivalent for both algorithms. Then, if there is no IMINT information and no interaction between targets, the performances of the algorithm with PCR5 are the same than without PCR5.

Despite of the PCR5 improvement on the target tracking, the difference of performances between the algorithms is not significant. If there is an interaction be-

Target	ARMSE	Track purity	Type purity
1	14.82	0.70	none
2	16.62	0.62	none
3	15.61	0.61	none
4	22.54	0.77	none
5	16.25	0.85	none
6	18.68	0.64	none
7	14.45	0.72	none
8	17.51	0.84	none
9	19.23	0.85	none
10	17.40	0.75	none

Table 1: Tracking results (VSIMMC without PCR5).

Target	ARMSE	Track purity	Type purity
1	14.37	0.78	0.64
2	15.77	0.66	0.62
3	15.60	0.61	0.59
4	21.10	0.81	0.81
5	15.88	0.94	0.55
6	18.68	0.64	0.02
7	14.22	0.76	0.76
8	17.38	0.87	0.87
9	19.20	0.85	0.05
10	17.17	0.83	0.46

Table 2: Tracking results (VSIMMC and PCR5).

tween IMINT and GMTI information, we can see a gain on the track length ratio or track purity of 10% with PCR5. This small difference is due to the good constrained state estimation. The estimated target states have a good precision because the target tracking is done by taking into account the road segments location and the good performances of the OOSM approach. So, it implies a substantial improvement of the target-to-track association. In addition, on Table 2, the type purity based on PCR5 is derived from the maximum of $BetP$ criterion. But $BetP$ is computed according the set C_{video} (12) and if the track receives only MTI reports the choice on the target type is arbitrary for the tracked vehicles of C_{MTI} (8). In fact, a tracked vehicle can be 6 elements of (12). So the probability $BetP$ on the 6 tracked vehicles of (12) is equivalent. The selection of the maximum of $BetP$ has no meaning because in such case and the maximum becomes arbitrary. This explains the bad track purity of targets 6 and 9.

6 Conclusion

In this paper, we have presented a new approach to improve VS IMMC SB-MHT by introducing the data fusion with several heterogeneous sensors. Starting from a centralized architecture, the MTI and IMINT reports are fused by taking into account the road network information and the OOSM algorithm for delayed measurements. The VS IMMC SB-MHT is enlarged by introducing in the data association process the type information defined in the STANAG 4607 and an IMINT attribute set. The estimation of the Target ID probability is done from the updated/current attribute mass of belief using the Proportional Conflict Redistribution rule no. 5 developed in DSmT framework and according to the Target Type Tracker (TTT) recently developed by the authors. The Target ID probability once obtained is then introduced in the track score computation in order to improve the likelihoods of each data association hypothesis of the SB-MHT. Our preliminary results show an improvement of the performances of the VS-IMMC SB-MHT when the type information is processed by our PCR5-based Target Type Tracker. In this work, we did not distinguish undelayed from delayed sensor reports in the TTT update. This problem is under investigations and offers new perspectives to find a solution for dealing efficiently with the time delay of the information type data and to improve performances. One simple solution would be to use a forgetting factor of the delayed type information but other solutions seem also possible to explore and need to be evaluated. Some works need also to be done to use the operational ontologic APP-6A for the heterogeneous type information. Actually, the frame of the IMINT type information is bigger than the one used in this paper and the IMINT type information can be given at different granularity levels. As a third perspective, we envisage to use both the type and contextual information in order to recognize the tracks losts in the terrain masks which represent the possible target occultations due to the terrain topography in real environments.

References

[1] T. Kirubarajan, and Y. Bar-Shalom, "Tracking evasive move-stop-move targets with an MTI radar using a VS-IMM estimator", *IEEE Tran. on AES*,Vol. 39, No. 3, pp. 1098–1103, Jul. 2003.

[2] M. Ulmke and W. Koch, "Road-map assisted ground moving target tracking", *IEEE Tran. on AES*, Vol. 42,No. 4, pp. 1264–1274, Oct. 2006.

[3] Y. Bar-Shalom and H. Chen, "IMM estimator with out-of-sequence measurements", *IEEE Tran. on AES*,Vol. 41, No. 1, pp. 90–98, Jan. 2005.

[4] E. Blash, B. Kahler, "Multiresolution EO/IR target tracking and identification",*Proc. of Fusion 2005*, Jul. 2005.

[5] B. Pannetier, J. Dezert and E. Pollard, "Improvement of multiple ground targets tracking with GMTI sensors and fusion identification attributes", *Aerospace Conference, 2008 IEEE*, Mar. 2008.

[6] B. Pannetier, V. Nimier and M. Rombaut, "Multiple ground target tracking with a GMTI sensor", *Proc. of MFI 2006*, Sept. 2006.

[7] NATO, "STANAG 4607 JAS (Edition 2) - NATO ground moving target indicator GMTI) format",*NSA0749(2007)-JAS/4607*, Aug. 2007.

[8] Y. Bar-Shalom and D. Blair, *Multitarget multisensor tracking : Applications and Advances*, Vol. III, Artech House, pp. 523, 2000.

[9] D.F. Bizup and D.E. Brown, "The over extended Kalman filter − Don't use it !", *Proc. of Fusion 2003*, Jul. 2003.

[10] S.S. Blackman and R. Popoli, *Design and analysis of modern tracking systems* , Artech House, 1999.

[11] F. Smarandache, J. Dezert, (Editors), *Advances and Applications of DSmT for Information Fusion (Collected Works)*, Vol. 2, American Research Press, Rehobooth, U.S.A, 2006.

[12] J. Dezert, A. Tchamova, F. Smarandache and P.Konstantinova, "Target Type Tracking with PCR5 and Dempster's rules: A Comparative Analysis", *Proc. of Fusion 2006*, Jul. 2006.

[13] Ph. Smets, "Data Fusion in the Transferable Belief Model", *Proc. of Fusion 2000*, Jul. 2000.

Threat assessment of a possible Vehicle-Born Improvised Explosive Device using DSmT

Jean Dezert
Florentin Smarandache

Originally published as Dezert J., Smarandache F., *Threat assessment of a possible Vehicle-Born Improvised Explosive Device using DSmT,* (presented during the forum on uncertainty at Fusion 2010 conference), Edinburgh, Scotland, UK, 26-29 July 2010, and printed with permission.

Abstract − *This paper presents the solution about the threat of a VBIED (Vehicle-Born Improvised Explosive Device) obtained with the DSmT (Dezert-Smarandache Theory). This problem has been proposed recently to the authors by Simon Maskell and John Lavery as a typical illustrative example to try to compare the different approaches for dealing with uncertainty for decision-making support. The purpose of this paper is to show in details how a solid justified solution can be obtained from DSmT approach and its fusion rules thanks to a proper modeling of the belief functions involved in this problem.*

Keywords: Security, Decison-making support, Information fusion, DSmT, Threat assessment.

1 The VBIED problem

- **Concern**: VBIED (Vehicle-Born Improvised Explosive Device) attack on an administrative building B
- **Prior information**: We consider an Individual A under surveillance due to previous unstable behavior who drives customized white Toyota (WT) vehicle.
- **Observation done at time** $t-10$ min: From a video sensor on road that leads to building B 10 min ago, one has observed a White Toyota 200m from the building B traveling in normal traffic flow toward building B. We consider the following two sources of information based on this video observation available at time $t-10$ min:
  - **Source 1**: An Analyst 1 with 10 years experience analyses the video and concludes that individual A is now probably near building B.
  - **Source 2**: An Automatic Number Plate Recognition (ANPR) system analyzing same video outputs 30% probability that the vehicle is individual A's white Toyota.
- **Observation done at time** $t-5$ min: From a video sensor on road 15km from building B 5 min ago one gets a video that indicates a white Toyota with some resemblance to individual A's white Toyota. We consider the following thrid source of information based on this video observation available at time $t-5$ min:
  - **Source 3**: An Analyst 2 (new in post) analyses this video and concludes that it is improbable that individual A is near building B.
- **Question 1**: Should building B be evacuated?
- **Question 2**: Is experience (Analyst 1) more valuable than physics (the ANPR system) combined with inexperience (Analyst 2)? How do we model that?

NOTE: Deception (e.g., individual A using different car, false number plates, etc.) and biasing (on the part of the analysts) are often a part of reality, but they are not part of this example.

2 Modeling the VBIED problem

Before applying DSmT fusion techniques to solve this VBIED problem it is important to model the problem in the framework of belief functions.

2.1 Marginal frames with their models

The marginal frames involved in this problem are:

- Frame related with individuals:

$$\Theta_1 = \{A = \text{Suspicious person}, \bar{A} = \text{not } A\}$$

- Frame related with the vehicle:

$$\Theta_2 = \{V = \text{White Toyota Vehicle}, \bar{V} = \text{not } V\}$$

- Frame related with the position of a driver of a car w.r.t the given building B:

$$\Theta_3 = \{B = \text{near building}, \bar{B} = \text{not } B\}$$

The underlying models of marginal frames are based on the following very reasonable assumptions:

- **Assumption 1**: We assume naturally $A \cap \bar{A} = \emptyset$ (avoiding Shrdinger's cat paradox). If working only with the frame of people Θ_3, the marginal bba's must be defined on the power-set

$$2^{\Theta_1} = \{\emptyset_1, A, \bar{A}, A \cup \bar{A}\}$$

- **Assumption 2**: We assume also that $V \cap \bar{V} = \emptyset$ so that the marginal bba (if needed) must be defined on the power-set

$$2^{\Theta_2} = \{\emptyset_2, V, \bar{V}, V \cup \bar{V}\}$$

- **Assumption 3**: We assume also that $B \cap \bar{B} = \emptyset$ so that the marginal bba (if needed) must be defined on the power-set

$$2^{\Theta_3} = \{\emptyset_3, B, \bar{B}, B \cup \bar{B}\}$$

This modeling is disputable since the notion of closeness/"near" is not clearly defined and we could prefer to work on

$$D^{\Theta_3} = \{\emptyset_3, B \cap \bar{B}, B, \bar{B}, B \cup \bar{B}\}$$

The emptyset elements have been indexed by the index of the frame they are referring to for notation convenience and avoiding confusion.

2.2 Joint frame and its model

Since we need to work with all aspects of available information, we need to define a common joint frame to express all what we have from different sources of information. The easiest way for defining the joint frame, denoted Θ, is to consider the classical Cartesian (cross) product space and to work with propositions (a Lindenbaum-Tarski algebra of propositions) since one has a correspondence between sets and propositions [5, 6], i.e.

$$\Theta = \Theta_1 \times \Theta_2 \times \Theta_3$$

which consists of the following 8 triplets elements

$$\begin{aligned}\Theta = \{&\theta_1 = (\bar{A}, \bar{V}, \bar{B}), \theta_2 = (A, \bar{V}, \bar{B}),\\ &\theta_3 = (\bar{A}, V, \bar{B}), \theta_4 = (A, V, \bar{B}),\\ &\theta_5 = (\bar{A}, \bar{V}, B), \theta_6 = (A, \bar{V}, B),\\ &\theta_7 = (\bar{A}, V, B), \theta_8 = (A, V, B)\}\end{aligned}$$

We define the union $\cup$, intersection $\cap$ as componentwise operators in the following way:

$$(x_1, x_2, x_3) \cup (y_1, y_2, y_3) \triangleq (x_1 \cup y_1, x_2 \cup y_2, x_3 \cup y_3)$$

$$(x_1, x_2, x_3) \cap (y_1, y_2, y_3) \triangleq (x_1 \cap y_1, x_2 \cap y_2, x_3 \cap y_3)$$

The complement $\bar{X}$ of X is defined in the usual way by

$$\bar{X} = \overline{(x_1, x_2, x_3)} \triangleq I_t \setminus \{X\}$$

where I_t is the total ignorance (i.e. the whole space of solutions) which corresponds to the maximal element defined by $I_t = (I_{t1}, I_{t2}, I_{t3})$, where I_{ti} is the maximal (ignorance) of Θ_i, $i = 1, 2, 3$. The minimum element (absolute empty proposition) is $\emptyset = (\emptyset_1, \emptyset_2, \emptyset_3)$, where $\emptyset_i$ is the minimum element (empty proposition) of Θ_i. We also define a relative minimum element in $S^{\Theta_1 \times \Theta_2 \times \Theta_3}$ as follows: $\emptyset_r = (x, y, z)$, where at least one of the components x, y, or z is a minimal element in its respective frame Θ_i. A general relative minimum element $\emptyset_{gr}$ is defined as the union/join of all relative minima (including the absolute minimum element). Similarly to the relative and general relative minimum we can define a relative maximum and a general relative maximum, where the empty set in the above definitions is replaced by the total ignorance. Whence the super-power set $(S^{\Theta}, \cap, \cup, {}^{-}, \emptyset, I_t)$ is equivalent to Lindenbaum-Tarski algebra of propositions.

For example, if we consider $\Theta_1 = \{x_1, x_2\}$ and $\Theta_2 = \{y_1, y_2\}$ satisfying both Shafer's model, then $\Theta = \Theta_1 \times \Theta_2 = \{(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)\}$, and one has:

$$\begin{aligned}\emptyset &= (\emptyset_1, \emptyset_2)\\ \emptyset_{r1} &= (\emptyset_1, y_1)\\ \emptyset_{r2} &= (\emptyset_1, y_2)\\ \emptyset_{r3} &= (\emptyset_1, y_1 \cup y_2)\\ \emptyset_{r4} &= (x_1, \emptyset_2)\\ \emptyset_{r5} &= (x_2, \emptyset_2)\\ \emptyset_{r6} &= (x_1 \cup x_2, \emptyset_2)\end{aligned}$$

and thus

$$\emptyset_{gr} = \emptyset \cup \emptyset_{r1} \cup \emptyset_{r2} \cup \ldots \cup \emptyset_{r6}$$

Based on definition of joint frame Θ with operations on its elements, we need to choose its underlying model (Shafer's, free or hybrid model) to define its fusion space where the bba's will be defined on. According to the definition of absolute and relative minimal elements, we then assume for the given VBIED problem that Θ satisfies Shafer's model, i.e. all (triplets) elements $\theta_i \in \Theta$ are exclusive, so that the bba's of sources will be defined on the classical power-set 2^{Θ}.

2.3 Supporting hypotheses for decision

In the VBIED problem the main question (Q1) is related with the security of people in the building B. The potential danger related with this building is of course $\theta_8 = (A, V, B)$ i.e. the presence of A in his/her car V near the building B. This is however and unfortunately not the only origin of the danger since the threat can also come from the possible presence of V (possible A's improvised explosive vehicle) parked near the building B even if A has left his/her car and is not himself/herself near the building. This second origin of danger is represented by $\theta_7 = (\bar{A}, V, B)$. There exists also a third origin of the danger represented by $\theta_6 = (A, \bar{V}, B)$ which reflects the possibility to have A near the building without V car. θ_6 is also dangerous for the building B since A can try to commit a suicidal terrorism attack as human bomb against the building. Therefore based on these three sources of potential danger, the most reasonable/prudent supporting hypothesis for decision-making is consider

$$\theta_6 \cup \theta_7 \cup \theta_8 = (A, \bar{V}, B) \cup (\bar{A}, V, B) \cup (A, V, B)$$

If we assume that the danger is mostly due to presence of A's vehicle containing possibly a high charge of explosive near the building B rather than the human bomb attack, then one can prefer to consider only the following hypothesis for decision-making support evaluation

$$\theta_7 \cup \theta_8 = (\bar{A}, V, B) \cup (A, V, B)$$

Finally if we are more *optimistic*, we can consider that the real danger occurs if and only if A drives V near the building B and therefore one could consider only the supporting hypothesis $\theta_8 = (A, V, B)$ for the danger in the decision-making support evaluation.

In the sequel, we adopt the worst scenario (we take the most prudent choice) and we consider all three origins of potential danger. Thus we will take $\theta_6 \cup \theta_7 \cup \theta_8$ as cautious/prudent supporting hypothesis for decision-making.

Thepropositions $\theta_6 \cup \theta_8 = (A, \bar{V}, B) \cup (A, V, B)$ and $\theta_6 \cup \theta_7 = (A, \bar{V}, B) \cup (\bar{A}, V, B)$ represent also a potential danger and could serve as decision-support hypotheses also, and their imprecise probabilities can be evaluate easily following analysis presented in the sequel. They have not been reported in this paper to keep it at a reasonable size.

2.4 Choice of bba's of sources

Let's define first the bba of each source without regard to what could be their reliability and importance in the fusion process. Reliability and importance will be examined in details in next section.

- **Bba related with source 0** (prior information): The prior information states that the suspect A drives a white Toyota, and nothing is state about the prior information with respect to his location, so that we must consider the bba's representing the prior information as

$$\begin{aligned} m_0(\theta_4 \cup \theta_8) &= m_0((A, V, \bar{B}) \cup (A, V, B)) \\ &= m_0((A, V, B \cup \bar{B})) \\ &= 1 \end{aligned}$$

- **Bba related with source 1** (Analyst 1 with 10 years experience): The source 1 reports that the suspect A is probably now near the building B. This source however doesn't report explicitly that the suspect A is still with its white Toyota car or not. So the fair way to model this report when working on Θ is to commit a high mass of belief to the element $\theta_6 \cup \theta_8$, that is

$$\begin{aligned} m_1(\theta_6 \cup \theta_8) &= m_1((A, \bar{V}, B) \cup (A, V, B)) \\ &= m_1((A, V \cup \bar{V}, B)) \\ &= 0.75 \end{aligned}$$

and to commit the uncommitted mass to I_t based on the principle of minimum of specificity, so that

$$m_1(\theta_6 \cup \theta_8) = 0.75 \quad \text{and} \quad m_1(I_t) = 0.25$$

- **Bba related with source 2** (ANPR system): The source 3 reports 30% probability that the vehicle is individual A's wite Toyota. Nothing is reported on the position information. The information provided by this source corresponds actually to incomplete probabilistic information. Indeed, when working on $\Theta_1 \times \Theta_2$, what we only know is that $P\{(A, V)\} = 0.3$ and $P\{(\bar{A}, V) \cup (A, \bar{V}) \cup (\bar{A}, \bar{V})\} = 0.7$ (from additivity axiom of probability theory) and thus the bba $m_2(.)$ we must choose on $\Theta_1 \times \Theta_2 \times \Theta_3$ has to be compatible with this incomplete probabilistic information, i.e. the projection $m'_2(.) \triangleq m_2^{\downarrow \Theta_1 \times \Theta_2}(.)$ of $m_2(.)$ on $\Theta_1 \times \Theta_2$ must satisfy the following constraints on belief and plausibility functions

$$Bel'((A, V)) = 0.3$$

$$Bel'((\bar{A}, V) \cup (A, \bar{V}) \cup (\bar{A}, \bar{V})) = 0.7$$

and also

$$Pl'((A, V)) = 0.3$$

$$Pl'((\bar{A}, V) \cup (A, \bar{V}) \cup (\bar{A}, \bar{V})) = 0.7$$

because belief and plausibility correspond to lower and upper bounds of probability measure [5]. So it is easy to verify that the following bba $m'_2(.)$

satisfy these constraints because the elements of the frame $\Theta_1 \times \Theta_2$ are exclusive:

$$m'_2((A,V)) = 0.3$$

$$m'_2((\bar{A},V) \cup (A,\bar{V}) \cup (\bar{A},\bar{V})) = 0.7$$

We can then extend $m'_2(.)$ into $\Theta_1 \times \Theta_2 \times \Theta_3$ using the minimum specificity principle (i.e. take the vacuous extension of $m'_2(.)$) to get the bba $m_2(.)$ that we need to solve the VBIED problem. That is $m_2(.) = m_2'^{\uparrow \Theta_1 \times \Theta_2 \times \Theta_3}(.)$ with

$$m_2((A,V,B \cup \bar{B})) = 0.3$$

$$m_2((\bar{A},V,B\cup\bar{B}) \cup (A,\bar{V},B\cup\bar{B}) \cup (\bar{A},\bar{V},B\cup\bar{B})) = 0.7$$

or equivalently

$$m_2(\theta_4 \cup \theta_8) = 0.3$$

$$m_2(\overline{\theta_4 \cup \theta_8}) = m_2(\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7) = 0.7$$

- **Bba related with source 3** (Analyst 3 with no experience): The source 3 reports that it is improbable that the suspect A is near the building B. This source however doesn't report explicitly that the suspect A is still with its white Toyota car or not. So the fair way to model this report when working on Θ is to commit a low mass of belief to the element $\theta_6 \cup \theta_8$, that is

$$\begin{aligned} m_3(\theta_6 \cup \theta_8) &= m_3((A,\bar{V},B) \cup (A,V,B)) \\ &= m_3((A, V \cup \bar{V}, B)) \\ &= 0.25 \end{aligned}$$

and to commit the uncommitted mass to I_t based on the principle of minimum of specificity, so that

$$m_3(\theta_6 \cup \theta_8) = 0.25 \quad \text{and} \quad m_3(I_t) = 0.75$$

2.5 Reliability of sources

Let's identify what is known about the reliability of sources and information:

- **Reliability of prior information**: it is (implicitly) supposed that the prior information is 100% reliable that is "Suspect A drives a white Toyota" which corresponds to the element $(A,V,B) \cup (A,V,\bar{B})$. So we can take the reliability factor of prior information as $\alpha_0 = 1$. If one considers the priori information highly reliable (but not totally reliable) then one could take $\alpha_0 = 0.9$ so that $m_0(.)$ would be

$$m_0(\theta_4 \cup \theta_8) = 0.9 \quad \text{and} \quad m_0(I_t) = 0.1$$

- **Reliability of source 1**: One knows that Analyst # 1 has 10 years experience, so we must consider him/her having a good reliability (say greater than 75%) or to be less precise we can just assign to him a qualitative reliability factor with minimal number of labels in $\{L_1 = \text{not good}, L_2 = \text{good}\}$. Here we should choose $\alpha_1 = L_2$. As first approximation, we can consider $\alpha_1 = 1$.

- **Reliability of source 2**: No information about the reliability of ANPR system is explicitly given. We may consider that if such device is used it is because it is also considered as a valuable tool and thus we assume it has a good reliability too, that is $\alpha_2 = 1$. If we want to be more prudent we should consider the reliability factor of this source as totally unknown and thus we should take it as very imprecise with $\alpha_2 = [0,1]$ (or qualitatively as $\alpha_2 = [L_0, L_3]$). If we are more optimistic and consider ANPR system as reliable enough, we could take α_2 a bit more precise with $\alpha_2 = [0.75, 1]$ (i.e. $\alpha_2 \geq 0.75$) or just qualitatively as $\alpha_2 = L_2$.

- **Reliability of source 3**: It is said explicitly that Analyst 2 is new in post, which means that Analyst 2 has no great experience and it can be inferred logically that it is less reliable than Analyst 1 so that we must choose $\alpha_3 < \alpha_1$. But we can also have a very young brillant analyst who perform very well too with respect to the older Analyst 1. So to be more cautious/prudent, we should also consider the case of unknown reliability factor α_3 by taking qualitatively $\alpha_3 = [L_0, L_3]$ or quantitatively by taking α_3 as a very imprecise value that is $\alpha_3 = [0,1]$.

2.6 Importance of sources

Not that much is explicitly said about the importance of the sources of information in the VBIED problem statement, but the fact that Analyst 1 has ten years experience and Analyst 2 is new in post, so that it seems logical to choose as importance factor $\beta_1 > \beta_3$. The importances discounting factors have been introduced and presented by the authors in [2, 7]. As a prudent attitude we could choose also $\beta_0 = [0,1] = [L_0, L_3]$ and $\beta_2 = [0,1] = [L_0, L_3]$ (vey imprecise values). If we consider that the prior information and the source 2 (ANPR) have the same importance, we could just take $\beta_0 = \beta_1 = 1$ to make derivations easier and adopt a more optimistic point of view[1].

3 Solution of VBIED problem

We apply PCR5 and PCR6 fusion rules developed originally in the DSmT framework to get the solution

[1] Of course the importance discounting factors can also be chosen approximatively from exogenous information upon the desiderata of the fusion system designer. This question is out of the scope of this paper.

of the VBIED problem. PCR5 has been developed by the authors in [6], Vol.2, and PCR6 is a variant of PCR5 proposed by Arnaud Martin and Christophe Osswald in [3]. Several codes for using PCR5 and PCR6 have been proposed in the literature for example in [3, 1, 7] and are available to the authors upon request.

Two cases are explored depending on the taking into account or not of the reliability and the importance of sources in the fusion process. To simplify the presentation of the results we denote the focal elements involved in this VBIED problem as:

$$f_1 \triangleq \theta_4 \cup \theta_8$$
$$f_2 \triangleq \theta_6 \cup \theta_8$$
$$f_3 \triangleq \theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7 = \overline{\theta_4 \cup \theta_8}$$
$$f_4 \triangleq I_t = \theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4 \cup \theta_5 \cup \theta_6 \cup \theta_7 \cup \theta_8$$

Only these focal elements are involved in inputs of the problem and we recall the two questions that we must answer:

Question 1 (Q1): Should building B be evacuated?
The question 1 must be answered by analyzing the level of belief and plausibility committed in the propositions supporting B through the fusion process.

Question 2 (Q2): Is experience (Analyst 1) more valuable than physics (the ANPR system) combined with inexperience (Analyst 2)? How do we model that?
The question 2 must be answered by analyzing and comparing the results of the fusion $m_1 \oplus m_3$ (or eventually $m_0 \oplus m_1 \oplus m_3$) with respect to m_2 only (resp. $m_0 \oplus m_2$).

3.1 Without reliability and importance

We provide here the solutions of the VBIED problem with direct PCR5 and PCR6 fusion of the sources for different qualitative inputs summarized in the tables below. We also present the result of DSmP probabilistic transformation [6] (Vol.3, Chap. 3) of resulting bba's to get and approximate probability measure of elements of Θ. No importance and reliability discounting has been applied since in this section, we consider that all sources have same importances and same reliabilities.

Example 1: We take the bba's described in section 2.3, that is

focal element	$m_0(.)$	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	1	0	0.3	0
$\theta_6 \cup \theta_8$	0	0.75	0	0.25
$\overline{\theta_4 \cup \theta_8}$	0	0	0.7	0
I_t	0	0.25	0	0.75

Table 1: Quantitative inputs of VBIED problem.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
$\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$	0.19741	0.16811
θ_8	0.24375	0.24375
$\theta_4 \cup \theta_8$	0.33826	0.29641
$\theta_6 \cup \theta_8$	0.11029	0.14587
I_t	0.11029	0.14587

Table 2: Results of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 1.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0333	0.0286
θ_2	0.0333	0.0286
θ_3	0.0333	0.0286
θ_4	0.0018	0.0018
θ_5	0.0333	0.0286
θ_6	0.0338	0.0292
θ_7	0.0333	0.0286
θ_8	0.7977	0.8260

Table 3: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 1.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0467	0.0463
θ_2	0.0467	0.0463
θ_3	0.0467	0.0463
θ_4	0.1829	0.1664
θ_5	0.0467	0.0463
θ_6	0.1018	0.1192
θ_7	0.0467	0.0463
θ_8	0.4818	0.4831

Table 4: $BetP$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 1.

From fusion result of Table 2, one gets for the danger supporting hypothesis $\theta_6 \cup \theta_7 \cup \theta_8$ (the worst scenario case)

- with PCR5: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.64596$
$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.35404, 1]$$
$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.64596]$$
- with PCR6: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.61038$
$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.38962, 1]$$
$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.61038]$$

where $\Delta(X) = Pl(X) - Bel(X)$ is the imprecision related to $P(X)$. It is worth to note that $\Delta(\bar{X}) = Pl(\bar{X}) - Bel(\bar{X}) = \Delta(X)$ because $Pl(\bar{X}) = 1 - Bel(X)$ and $Bel(\bar{X}) = 1 - Pl(X)$.

If we consider only $\theta_7 \cup \theta_8 = (\bar{A}, V, B) \cup (A, V, B)$ as danger supporting hypothesis then from the fusion result of Table 2, one gets

- with PCR5: $\Delta(\theta_7 \cup \theta_8) = 0.75625$
$$P(\theta_7 \cup \theta_8) \in [0.24375, 1]$$
$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.75625]$$

- with PCR6: $\Delta(\theta_7 \cup \theta_8) = 0.75625$

$$P(\theta_7 \cup \theta_8) \in [0.24375, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.75625]$$

If we are more *optimistic* and we consider only the danger supporting hypothesis θ_8, then one gets

- with PCR5: $\Delta(\theta_8) = 0.55884$

$$P(\theta_8) \in [0.24375, 0.80259]$$

$$P(\bar{\theta}_8) \in [0.19741, 0.75625]$$

- with PCR6: $\Delta(\theta_8) = \Delta(\bar{\theta}_8) = 0.58814$

$$P(\theta_8) \in [0.24375, 0.83189]$$

$$P(\bar{\theta}_8) \in [0.16811, 0.75625]$$

If one approximates the bba's into probabilistic measures with DSmP transformation[2], one gets results with $\epsilon = 0.001$ presented in Table 3. One gets the higher probability on θ_8 with respect to other alternatives and also $DSmP(\theta_6 \cup \theta_7 \cup \theta_8) = 0.8648$. If one prefers to use the pignistic[3] probability transformation [8], one gets the results given in Table 4. One sees clearly that PIC[4] of DSmP is higher that PIC of BetP which makes decision easier to take with DSmP than with BetP in favor of $\theta_6 \cup \theta_7 \cup \theta_8$, or $\theta_7 \cup \theta_8$, or θ_8.

- **Answer to Q1:** One sees that the result provided by PCR6 and PCR5 are very close and do not change fundamentally the final decision to take. Based on these very imprecise results, it is very difficult to take the right decision without decision-making error because the sources of information are highly uncertain and conflicting, but the analysis of lower and upper bounds shows that the most reasonable answer to the question based either on max of credibility or max of plausibility is to evacuate the building B since $Bel(\theta_6 \cup \theta_7 \cup \theta_8) > Bel(\overline{\theta_6 \cup \theta_7 \cup \theta_8})$ and also $Pl(\theta_6 \cup \theta_7 \cup \theta_8) > Pl(\overline{\theta_6 \cup \theta_7 \cup \theta_8})$. The same conclusion is drawn when considering the element $\theta_7 \cup \theta_8$ or θ_8 alone. The same conclusion also is drawn (more easier) based on DSmP or on BetP values. In summary, the answer to Q1 is: **Evacuation of the building** B.

[2]DSmP transformation has been introduced and justified in details by the authors in the book [6] (Vol.3, Chap. 3) freely downloadable from the web with many examples, and therefore it will not be presented here.

[3]BetP is the most used transformation to approximate a mass of belief into a subjective probability measure. It has been proposed by Philippe Smets in nineties.

[4]The PIC (probabilistic information content) criteria has been introduced by John Sudano in [9] and is noting but the dual of normalized Shannon entropy. PIC is in $[0, 1]$ and $PIC = 1$ if the probability measure assigns a probability one only on a particular singleton of the frame, and $PIC = 0$ if all elements of the frame are equi-probable.

In order to answer to the second question (Q2), let's compute the fusion results of the fusion $m_0 \oplus m_2$ and $m_0 \oplus m_1 \oplus m_3$ using inputs given in Table 1. The fusion results with corresponding DSmP and BetP are given in the Tables 5-6.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
$\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$	0.28824	0.28824
$\theta_4 \cup \theta_8$	0.71176	0.71176

Table 5: Result of $m_0 \oplus m_2$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0480	0.0480
θ_2	0.0480	0.0480
θ_3	0.0480	0.0480
θ_4	0.3560	0.3560
θ_5	0.0480	0.0480
θ_6	0.0480	0.0480
θ_7	0.0480	0.0480
θ_8	0.3560	0.3560

Table 6: $DSmP_\epsilon$ of $m_0 \oplus m_2$.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0480	0.0480
θ_2	0.0480	0.0480
θ_3	0.0480	0.0480
θ_4	0.3560	0.3560
θ_5	0.0480	0.0480
θ_6	0.0480	0.0480
θ_7	0.0480	0.0480
θ_8	0.3560	0.3560

Table 7: $BetP$ of $m_0 \oplus m_2$.

Based on $m_0 \oplus m_2$ fusion result, one gets a total imprecision $\Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8) = 1$ when considering $\theta_6 \cup \theta_7 \cup \theta_8$ or $\theta_7 \cup \theta_8$ since

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 1]$$

and

$$P(\theta_7 \cup \theta_8) \in [0, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 1]$$

Even when considering only the danger supporting hypothesis θ_8, one still gets a quite large imprecision on $P(\theta_8)$ since $\Delta_{02}(\theta_8) = 0.71176$ with

$$P(\theta_8) \in [0, 0.71176]$$

$$P(\bar{\theta}_8) \in [0.28824, 1]$$

Based on max of Bel or max of Pl criteria, one sees that it is not possible to take any rational decision from $\theta_6 \cup \theta_7 \cup \theta_8$ nor $\theta_7 \cup \theta_8$ because of the full imprecision

range of $P(\theta_6 \cup \theta_7 \cup \theta_8)$ or $P(\theta_7 \cup \theta_8)$. The decision using $m_0 \oplus m_2$ (i.e. with prior information m_0 and ANPR system m_2) based only on supporting hypothesis θ_8 should be to NOT evacuate the building B. Same decision would be taken based on DSmP or BetP values.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_8	0.8125	0.8125
$\theta_4 \cup \theta_8$	0.1875	0.1875

Table 8: Result of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0	0
θ_2	0	0
θ_3	0	0
θ_4	0.0002	0.0002
θ_5	0	0
θ_6	0	0
θ_7	0	0
θ_8	0.9998	0.9998

Table 9: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0	0
θ_2	0	0
θ_3	0	0
θ_4	0.09375	0.09375
θ_5	0	0
θ_6	0	0
θ_7	0	0
θ_8	0.90625	0.90625

Table 10: $BetP$ of $m_0 \oplus m_1 \oplus m_3$.

Based on $m_0 \oplus m_1 \oplus m_3$ fusion result, one gets $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.1875$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.8125, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.1875]$$

but also

$$P(\theta_7 \cup \theta_8) \in [0.8125, 1], \qquad P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.1875]$$

and

$$P(\theta_8) \in [0.8125, 1], \qquad P(\bar{\theta}_8) \in [0, 0.1875]$$

Based on max of Bel or max of Pl criteria, the decision using $m_0 \oplus m_1 \oplus m_3$ (i.e. with prior information m_0 and both analysts) is to evacuate the building B. The same decision is taken based on DSmP or BetP values. It is worth to note that the precision on the result obtained with $m_0 \oplus m_1 \oplus m_3$ is much better than with $m_0 \oplus m_2$ since $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) < \Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8)$, or $\Delta_{013}(\theta_7 \cup \theta_8) < \Delta_{02}(\theta_7 \cup \theta_8)$. Moreover it is easy to verify that $m_0 \oplus m_1 \oplus m_3$ fusion system is more informative than $m_0 \oplus m_2$ fusion system because Shannon entropy of DSmP of $m_0 \oplus m_2$ is much bigger than Shannon entropy of DSmP of $m_0 \oplus m_1 \oplus m_3$.

- **Answer to Q2:** Since the information obtained by the fusion $m_0 \oplus m_2$ is less informative and less precise than the information obtained with the fusion $m_0 \oplus m_1 \oplus m_3$, it is better to choose and to trust the fusion system $m_0 \oplus m_1 \oplus m_3$ rather than $m_0 \oplus m_2$. Based on this choice, the final decision will be to evacuate the building B which is consistent with answer to question Q1.

Example 2: Let's modify a bit the previous Table 1 and take higher belief for sources 1 and 3 as

focal element	$m_0(.)$	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	1	0	0.3	0
$\theta_6 \cup \theta_8$	0	0.9	0	0.1
$\overline{\theta_4 \cup \theta_8}$	0	0	0.7	0
I_t	0	0.1	0	0.9

Table 11: Quantitative inputs of VBIED problem.

The results of the fusion $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ using PCR5 and PCR6 and the corresponding DSmP values are given in tables 12-13.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
$\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$	0.16525	0.14865
θ_8	0.27300	0.27300
$\theta_4 \cup \theta_8$	0.26307	0.23935
$\theta_6 \cup \theta_8$	0.14934	0.16950
I_t	0.14934	0.16950

Table 12: Results of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 11.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0281	0.0254
θ_2	0.0281	0.0254
θ_3	0.0281	0.0254
θ_4	0.0015	0.0015
θ_5	0.0281	0.0254
θ_6	0.0286	0.0260
θ_7	0.0281	0.0254
θ_8	0.8294	0.8455

Table 13: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 11.

From fusion result of Table 12, one gets when considering $\theta_6 \cup \theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.57766$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.42234, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 57766]$$

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0462	0.0460
θ_2	0.0462	0.0460
θ_3	0.0462	0.0460
θ_4	0.1502	0.1409
θ_5	0.0462	0.0460
θ_6	0.1209	0.1307
θ_7	0.0462	0.0460
θ_8	0.4979	0.4986

Table 14: $BetP$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 11.

- with PCR6: $\Delta(\theta_7 \cup \theta_8) = 0.5575$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.4425, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.5575]$$

and when considering $\theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_7 \cup \theta_8) = 0.7270$

$$P(\theta_7 \cup \theta_8) \in [0.27300, 1], \quad P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.7270]$$

- with PCR6: $\Delta(\theta_7 \cup \theta_8) = 0.7270$

$$P(\theta_7 \cup \theta_8) \in [0.27300, 1], \quad P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.7270]$$

and when considering θ_8 only

- with PCR5 or PCR6: $\Delta(\theta_8) = 0.56175$

$$P(\theta_8) \in [0.27300, 0.83475]$$

$$P(\bar{\theta}_8) \in [0.16525, 0.7270]$$

- with PCR6: $\Delta(\theta_8) = 0.57835$

$$P(\theta_8) \in [0.27300, 0.85135]$$

$$P(\bar{\theta}_8) \in [0.14865, 0.7270]$$

One gets also the following $DSmP_{\epsilon=0.001}$ values

$$DSmP_{\epsilon,PCR5}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.8861$$
$$DSmP_{\epsilon,PCR6}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.8869$$
$$DSmP_{\epsilon,PCR5}(\theta_7 \cup \theta_8) = 0.8575$$
$$DSmP_{\epsilon,PCR6}(\theta_7 \cup \theta_8) = 0.8709$$
$$DSmP_{\epsilon,PCR5}(\theta_8) = 0.8294$$
$$DSmP_{\epsilon,PCR6}(\theta_8) = 0.8455$$

- **Answer to Q1:** Using an analysis similar to the one done for Example 1, based on max of credibility or max of plausibility criteria, or by considering the DSmP or BetP values of $\theta_6 \cup \theta_7 \cup \theta_8$, or $\theta_7 \cup \theta_8$, or θ_8 the decision to take is: **Evacuate the building** B.

In order to answer to the second question (Q2) for this Example 2, let's compute the fusion results of the fusion $m_0 \oplus m_2$ and $m_0 \oplus m_1 \oplus m_3$ using inputs given in Table 11. Since the inputs m_0 and m_2 are the same as those in Example 1, the $m_0 \oplus m_2$ fusion results with corresponding DSmP are those already given in Tables 5-7. Only the fusion $m_0 \oplus m_1 \oplus m_3$ must be derived with the new bba's m_1 and m_3 chosen for this Example 2. The $m_0 \oplus m_1 \oplus m_3$ fusion results obtained with PCR5 and PCR6, and the corresponding DSmP and BetP values are shown in Tables 15-17. According to these results, one gets with the PCR5 or PCR6 fusion $m_0 \oplus m_1 \oplus m_3$: $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) = \Delta_{013}(\theta_7 \cup \theta_8) = \Delta_{013}(\theta_8) = 0.09$ and

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.91, 1], \quad P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.09]$$

$$P(\theta_7 \cup \theta_8) \in [0.91, 1], \quad P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.09]$$

$$P(\theta_8) \in [0.91, 1], \quad P(\bar{\theta}_8) \in [0, 0.09]$$

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_8	0.91	0.91
$\theta_4 \cup \theta_8$	0.09	0.09

Table 15: Result of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0	0
θ_2	0	0
θ_3	0	0
θ_4	0.0001	0.0001
θ_5	0	0
θ_6	0	0
θ_7	0	0
θ_8	0.9999	0.9999

Table 16: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0	0
θ_2	0	0
θ_3	0	0
θ_4	0.045	0.045
θ_5	0	0
θ_6	0	0
θ_7	0	0
θ_8	0.955	0.955

Table 17: $BetP$ of $m_0 \oplus m_1 \oplus m_3$.

Based on max of Bel or max of Pl criteria, the decision using $m_0 \oplus m_1 \oplus m_3$ (i.e. with prior information m_0 and both analysts) is to evacuate the building B. Same decision is taken based on DSmP or BetP values. It is worth to note that the precision on the result obtained with $m_0 \oplus m_1 \oplus m_3$ is much better than with $m_0 \oplus m_2$ since $\Delta_{013}(\theta_8) < \Delta_{02}(\theta_8)$, or $\Delta_{013}(\theta_7 \cup \theta_8) <$

$\Delta_{02}(\theta_7 \cup \theta_8)$, or $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) < \Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8)$. Moreover it is easy to verify that $m_0 \oplus m_1 \oplus m_3$ fusion system is more informative than $m_0 \oplus m_2$ fusion system because Shannon entropy of DSmP of $m_0 \oplus m_2$ is much bigger than Shannon entropy of DSmP of $m_0 \oplus m_1 \oplus m_3$. Same remark holds with BetP transformation.

- **Answer to Q2:** Since the information obtained by the fusion $m_0 \oplus m_2$ is less informative and less precise than the information obtained with the fusion $m_0 \oplus m_1 \oplus m_3$, it is better to choose and to trust the fusion system $m_0 \oplus m_1 \oplus m_3$ rather than $m_0 \oplus m_2$. Based on this choice, the final decision will be to evacuate the building B which is consistent with the answer of the question Q1.

3.2 Impact of prior information

To see the impact of the quality/reliability of prior information on the result, let's modify the input $m_0(.)$ in previous Tables 1 and 11 and consider now a very uncertain prior source.

Example 3: We consider the very uncertain prior source of information $m_0(\theta_4 \cup \theta_8) = 0.1$ and $m_0(I_t) = 0.9$. The results for the modified inputs Table 18 ate given in Tables 19 and 20.

focal element	$m_0(.)$	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	0.1	0	0.3	0
$\theta_6 \cup \theta_8$	0	0.75	0	0.25
$\overline{\theta_4 \cup \theta_8}$	0	0	0.7	0
I_t	0.9	0.25	0	0.75

Table 18: Quantitative inputs of VBIED problem.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_6	0.511870	0.511870
$\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$	0.151070	0.142670
θ_8	0.243750	0.243750
$\theta_4 \cup \theta_8$	0.060957	0.059757
$\theta_6 \cup \theta_8$	0.016173	0.020973
I_t	0.016173	0.020973

Table 19: Result of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 18.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0003	0.0003
θ_2	0.0003	0.0003
θ_3	0.0003	0.0003
θ_4	0.0003	0.0003
θ_5	0.0003	0.0003
θ_6	0.6833	0.6815
θ_7	0.0003	0.0003
θ_8	0.3149	0.3168

Table 20: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 18.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0272	0.0264
θ_2	0.0272	0.0264
θ_3	0.0272	0.0264
θ_4	0.0325	0.0325
θ_5	0.0272	0.0264
θ_6	0.5472	0.5488
θ_7	0.0272	0.0264
θ_8	0.2843	0.2867

Table 21: $BetP$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 18.

From the fusion result of Table 19, one gets when considering $\theta_6 \cup \theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.221377$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.778623, 1]$$
$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.221377]$$

- with PCR6: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.21465$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.78535, 1]$$
$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.21465]$$

and when considering $\theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_7 \cup \theta_8) = 0.24438$

$$P(\theta_7 \cup \theta_8) \in [0.24375, 0.48813]$$
$$P(\overline{\theta_7 \cup \theta_8}) \in [0.51187, 0.75625]$$

- with PCR6: $\Delta(\theta_7 \cup \theta_8) = 0.24438$

$$P(\theta_7 \cup \theta_8) \in [0.24375, 0.48813]$$
$$P(\overline{\theta_7 \cup \theta_8}) \in [0.51187, 0.75625]$$

and when considering θ_8 only, one has

- with PCR5: $\Delta(\theta_8) = 0.09331$

$$P(\theta_8) \in [0.24375, 0.33706]$$
$$P(\bar{\theta}_8) \in [0.66294, 0.75625]$$

- with PCR6: $\Delta(\theta_8) = 0.10171$

$$P(\theta_8) \in [0.24375, 0.34546]$$
$$P(\bar{\theta}_8) \in [0.65454, 0.75625]$$

Using DSmP transformation, one gets a low probability in θ_8 or in $\theta_7 \cup \theta_8$ because

$$DSmP_{\epsilon,PCR5}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.9985$$
$$DSmP_{\epsilon,PCR6}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.9986$$
$$DSmP_{\epsilon,PCR5}(\theta_7 \cup \theta_8) = 0.3152$$
$$DSmP_{\epsilon,PCR6}(\theta_7 \cup \theta_8) = 0.3171$$
$$DSmP_{\epsilon,PCR5}(\theta_8) = 0.3149$$
$$DSmP_{\epsilon,PCR6}(\theta_8) = 0.3168$$

- **Answer to Q1:** The analysis of these results are very interesting since one sees that the element of the frame Θ having the highest DSmP (or BetP) is $\theta_6 = (A, \bar{V}, B)$ and it has a very strong impact on the final decision. Because if one considers only θ_8 or $\theta_7 \cup \theta_8$ has decision-support hypotheses, one sees that the decision to take is to NOT evacuate the building B since one gets a low probability in θ_8 or in $\theta_7 \cup \theta_8$. Whereas if we include also θ_6 in the decision-support hypothesis, then the final decision will be the opposite since $DSmP(\theta_6 \cup \theta_7 \cup \theta_8)$ is very close to one with PCR5 or with PCR6. The same behavior occurs with BetP. So there is a strong impact of prior information on the final decision since without strong prior information supporting $\theta_4 \cup \theta_8$ we have to conclude either to the non evacuation of building B based on the max of credibility, the max of plausibility or the max of DSmP using θ_8 or $\theta_7 \cup \theta_8$ for decision-making, or to the evacuation of the building if a more prudent strategy is used based on $\theta_6 \cup \theta_7 \cup \theta_8$ decision-support hypothesis.

Let's examine the results of fusion systems $m_0 \oplus m_2$ and $m_0 \oplus m_1 \oplus m_3$ given in Tables 22-27.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
$\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$	0.69125	0.69125
$\theta_4 \cup \theta_8$	0.30875	0.30875

Table 22: Result of $m_0 \oplus m_2$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.1152	0.1152
θ_2	0.1152	0.1152
θ_3	0.1152	0.1152
θ_4	0.1544	0.1544
θ_5	0.1152	0.1152
θ_6	0.1152	0.1152
θ_7	0.1152	0.1152
θ_8	0.1544	0.1544

Table 23: $DSmP_\epsilon$ of $m_0 \oplus m_2$.

Singletons	$BetP_{\epsilon,PCR5}(.)$	$BetP_{\epsilon,PCR6}(.)$
θ_1	0.1152	0.1152
θ_2	0.1152	0.1152
θ_3	0.1152	0.1152
θ_4	0.1544	0.1544
θ_5	0.1152	0.1152
θ_6	0.1152	0.1152
θ_7	0.1152	0.1152
θ_8	0.1544	0.1544

Table 24: $BetP$ of $m_0 \oplus m_2$.

Based on $m_0 \oplus m_2$ fusion result, one gets a large imprecision on evaluation of probabilities of decision-support hypotheses since for $\theta_6 \cup \theta_7 \cup \theta_8$, one has

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_8	0.08125	0.08125
$\theta_4 \cup \theta_8$	0.01875	0.01875
$\theta_6 \cup \theta_8$	0.73125	0.73125
I_t	0.16875	0.16875

Table 25: Result of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0019	0.0019
θ_2	0.0019	0.0019
θ_3	0.0019	0.0019
θ_4	0.0021	0.0021
θ_5	0.0019	0.0019
θ_6	0.0107	0.0107
θ_7	0.0019	0.0019
θ_8	0.9778	0.9778

Table 26: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$BetP_{\epsilon,PCR5}(.)$	$BetP_{\epsilon,PCR6}(.)$
θ_1	0.0211	0.0211
θ_2	0.0211	0.0211
θ_3	0.0211	0.0211
θ_4	0.0305	0.0305
θ_5	0.0211	0.0211
θ_6	0.3867	0.3867
θ_7	0.0211	0.0211
θ_8	0.4773	0.4773

Table 27: $BetP$ of $m_0 \oplus m_1 \oplus m_3$.

$\Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8) = 1$ and

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0,1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0,1]$$

for $\theta_7 \cup \theta_8$, one has also $\Delta_{02}(\theta_7 \cup \theta_8) = 1$ with

$$P(\theta_7 \cup \theta_8) \in [0,1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0,1]$$

and for θ_8, one gets $\Delta_{02}(\theta_8) = 0.30875$ with

$$P(\theta_8) \in [0, 0.30875]$$

$$P(\bar{\theta}_8) \in [0.69125, 1]$$

One sees that it is impossible to take a decision when considering only $\theta_6 \cup \theta_7 \cup \theta_8$ or $\theta_7 \cup \theta_8$ because of full imprecision of the corresponding probabilities. However, based on max of Bel or max of Pl criteria on θ_8 the decision using $m_0 \oplus m_2$ (i.e. with uncertain prior information m_0 and ANPR system m_2) is to NOT evacuate the building B. According to Tables 23-24, one sees also an ambiguity in decision-making between θ_8 and θ_4 since they have the same DSmP (or BetP) values.

Based on $m_0 \oplus m_1 \oplus m_3$ fusion results given in Tables 25-27, one gets for $\theta_6 \cup \theta_7 \cup \theta_8$ the imprecision $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.1875$ with

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.8125, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.1875]$$

for $\theta_7 \cup \theta_8$, one gets $\Delta_{013}(\theta_7 \cup \theta_8) = 0.91875$ with

$$P(\theta_7 \cup \theta_8) \in [0.08125, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.91875]$$

and for θ_8, one gets $\Delta_{013}(\theta_8) = 0.91875$ with

$$P(\theta_8) \in [0.08125, 1]$$

$$P(\bar{\theta}_8) \in [0, 0.91875]$$

Based on max of Bel or max of Pl criteria, the decision using $m_0 \oplus m_1 \oplus m_3$ is the evacuation of the building B. Same decision is drawn when using DSmP or BetP results according to Tables 26 and 27. With this uncertain prior information, it is worth to note that the precision on the result obtained with $m_0 \oplus m_1 \oplus m_3$ is better than with $m_0 \oplus m_2$ when considering (in cautious strategy) the decision-support hypotheses $\theta_6 \cup \theta_7 \cup \theta_8$ or $\theta_7 \cup \theta_8$ since $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) < \Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8)$, or $\Delta_{013}(\theta_7 \cup \theta_8) < \Delta_{02}(\theta_7 \cup \theta_8)$. However, if a more optimistic/risky strategy is used when considering only θ_8 as decision-support hypothesis, it is preferable to choose the subsystem $m_0 \oplus m_2$ because $\Delta_{02}(\theta_8) < \Delta_{013}(\theta_8)$. However, one sees that globally $m_0 \oplus m_1 \oplus m_3$ fusion system is more informative than $m_0 \oplus m_2$ fusion system because Shannon entropy of DSmP of $m_0 \oplus m_2$ is much bigger than Shannon entropy of DSmP of $m_0 \oplus m_1 \oplus m_3$.

- **Answer to Q2:** The answer of question Q2 is not easy because it depends both on the criterion (precision or PIC) and on the decision-support hypothesis we choose. Based on precision criterion and taking the optimistic point of view using only θ_8, it is better to trust $m_0 \oplus m_2$ fusion system since $\Delta_{02}(\theta_8) = \Delta_{02}(\bar{\theta}_8) = 0.30875$ whereas $\Delta_{013}(\theta_8) = \Delta_{013}(\bar{\theta}_8) = 0.9187$. In such case, one should NOT evacuate the building B. If we consider that is better to trust result of $m_0 \oplus m_1 \oplus m_3$ fusion system because it is more informative than $m_0 \oplus m_2$ then the decision should be to evacuate the building B. If we take a more prudent point of view in considering as decision-support hypotheses either $\theta_6 \cup \theta_7 \cup \theta_8$ or $\theta_7 \cup \theta_8$, then the final decision taken according to the (most precise and informative) subsystem $m_0 \oplus m_1 \oplus m_3$ is to evacuate the building B.

  So the main open question is what solution to choose for selecting either $m_0 \oplus m_2$ or $m_0 \oplus m_1 \oplus m_3$ fusion system ? In authors opinion, in such case it seems better to base our choice on the precision level of information one has really in hands (rather than the PIC value which is always related to some ad-hoc probabilistic transformation) and in adopting the most prudent strategy. Therefore for this example, the final decision must be done according to $m_0 \oplus m_1 \oplus m_3$, i.e. evacuate the building B.

Example 4: Let's modify a bit the previous input Table 18 and take higher belief for sources 1 and 3 as

focal element	$m_0(.)$	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	0.1	0	0.3	0
$\theta_6 \cup \theta_8$	0	0.9	0	0.1
$\overline{\theta_4 \cup \theta_8}$	0	0	0.7	0
I_t	0.9	0.1	0	0.9

Table 28: Quantitative inputs of VBIED problem.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_6	0.573300	0.573300
$\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$	0.082365	0.077355
θ_8	0.273000	0.273000
$\theta_4 \cup \theta_8$	0.030666	0.029951
$\theta_6 \cup \theta_8$	0.020334	0.023197
I_t	0.020334	0.023197

Table 29: Result of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 28.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0002	0.0002
θ_2	0.0002	0.0002
θ_3	0.0002	0.0002
θ_4	0.0001	0.0001
θ_5	0.0002	0.0002
θ_6	0.6824	0.6813
θ_7	0.0002	0.0002
θ_8	0.3166	0.3178

Table 30: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 28.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0163	0.0158
θ_2	0.0163	0.0158
θ_3	0.0163	0.0158
θ_4	0.0179	0.0179
θ_5	0.0163	0.0158
θ_6	0.5997	0.6007
θ_7	0.0163	0.0158
θ_8	0.3010	0.3025

Table 31: $BetP$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 28.

Therefore, one gets when considering $\theta_6 \cup \theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.133366$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.866634, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.133366]$$

- with PCR6: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.130503$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.869497, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.130503]$$

and when considering $\theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_7 \cup \theta_8) = 0.1537$

$$P(\theta_7 \cup \theta_8) \in [0.2730, 0.4267]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0.5733, 0.7270]$$

- with PCR6: $\Delta(\theta_7 \cup \theta_8) = 0.1537$

$$P(\theta_7 \cup \theta_8) \in [0.2730, 0.4267]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0.5733, 0.7270]$$

and when considering θ_8 only

- with PCR5: $\Delta(\theta_8) = 0.071335$

$$P(\theta_8) \in [0.2730, 0.344335]$$

$$P(\bar{\theta}_8) \in [0.655665, 0.7270]$$

- with PCR6: $\Delta(\theta_8) = 0.076345$

$$P(\theta_8) \in [0.2730, 0.349345]$$

$$P(\bar{\theta}_8) \in [0.650655, 0.7270]$$

Based on DSmP transformation, one gets a pretty low probability on θ_8 and on $\theta_7 \cup \theta_8$, but a very high probability on the most prudent decision-support fypothesis $\theta_6 \cup \theta_7 \cup \theta_8$ because

$$\begin{aligned} DSmP_{\epsilon,PCR5}(\theta_6 \cup \theta_7 \cup \theta_8) &= 0.9992 \\ DSmP_{\epsilon,PCR6}(\theta_6 \cup \theta_7 \cup \theta_8) &= 0.9993 \\ DSmP_{\epsilon,PCR5}(\theta_7 \cup \theta_8) &= 0.3168 \\ DSmP_{\epsilon,PCR6}(\theta_7 \cup \theta_8) &= 0.3180 \\ DSmP_{\epsilon,PCR5}(\theta_8) &= 0.3166 \\ DSmP_{\epsilon,PCR6}(\theta_8) &= 0.3178 \end{aligned}$$

- **Answer to Q1:** Based on these results, one sees that the decision based either on the max of credibility, the max of plausibility or the max of DSmP considering both cases θ_8 or $\theta_7 \cup \theta_8$ is to: NOT Evacuate the building B, whereas the most prudent/cautious strategy suggests the opposite, i.e. the evacuation of the building B.

Let's examine the results of fusion systems $m_0 \oplus m_2$ and $m_0 \oplus m_1 \oplus m_3$ corresponding to the input Table 28. Naturally, one gets same results for the fusion $m_0 \oplus m_2$ as in Example 3 and for the fusion $m_0 \oplus m_1 \oplus m_3$ one gets:

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_8	0.091	0.091
$\theta_4 \cup \theta_8$	0.009	0.009
$\theta_6 \cup \theta_8$	0.819	0.819
I_t	0.081	0.081

Table 32: Result of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0008	0.0008
θ_2	0.0008	0.0008
θ_3	0.0008	0.0008
θ_4	0.0009	0.0009
θ_5	0.0008	0.0008
θ_6	0.0096	0.0096
θ_7	0.0008	0.0008
θ_8	0.9854	0.9854

Table 33: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0101	0.0101
θ_2	0.0101	0.0101
θ_3	0.0101	0.0101
θ_4	0.0146	0.0146
θ_5	0.0101	0.0101
θ_6	0.4196	0.4196
θ_7	0.0101	0.0101
θ_8	0.5151	0.5151

Table 34: $BetP$ of $m_0 \oplus m_1 \oplus m_3$.

As in Example 3, based on $m_0 \oplus m_2$ fusion result, one gets a large imprecision on evaluation of probabilities of decision-support hypotheses since for $\theta_6 \cup \theta_7 \cup \theta_8$, one has $\Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8) = 1$ and

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 1]$$

for $\theta_7 \cup \theta_8$, one has also $\Delta_{02}(\theta_7 \cup \theta_8) = 1$ with

$$P(\theta_7 \cup \theta_8) \in [0, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 1]$$

and for θ_8, one gets $\Delta_{02}(\theta_8) = 0.30875$ with

$$P(\theta_8) \in [0, 0.30875]$$

$$P(\bar{\theta}_8) \in [0.69125, 1]$$

One sees that it is impossible to take a decision when considering only $\theta_6 \cup \theta_7 \cup \theta_8$ or $\theta_7 \cup \theta_8$ because of full imprecision of the corresponding probabilities. Based on max of Bel or max of Pl criteria on θ_8 the decision using $m_0 \oplus m_2$ is to NOT evacuate the building B. According to Tables 23-24, an ambiguity appears in decision-making between θ_8 and θ_4 since they have the same DSmP (or BetP) values.

Based on $m_0 \oplus m_1 \oplus m_3$ fusion result, one gets $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.09$ for $\theta_6 \cup \theta_7 \cup \theta_8$ with

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.91, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.09]$$

and $\Delta_{013}(\theta_7 \cup \theta_8) = 0.9090$ for $\theta_7 \cup \theta_8$ with

$$P(\theta_7 \cup \theta_8) \in [0.091, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.9090]$$

and $\Delta_{013}(\theta_8) = 0.9090$ for only θ_8 with

$$P(\theta_8) \in [0.091, 1]$$

$$P(\bar{\theta}_8) \in [0, 0.9090]$$

Based on max of Bel or max of Pl criteria on either θ_8, $\theta_7 \cup \theta_8$ or $\theta_6 \cup \theta_7 \cup \theta_8$ the decision using $m_0 \oplus m_1 \oplus m_3$ must be the evacuation of the building B. Same decision is drawn using DSmP or BetP results according to Tables 33 and 34.

- **Answer to Q2:** Similar remarks and conclusions to those given in Example 3 held also for Example 4, i.e. it is better to adopt the most prudent strategy (i.e. to consider $\theta_6 \cup \theta_7 \cup \theta_8$ as decision-support hypothesis) and to trust the most precise fusion system with respect this hypothesis, which is in this example the subsystem $m_0 \oplus m_1 \oplus m_3$. Based only on $m_0 \oplus m_1 \oplus m_3$ the final decision will be to evacuate the building B when one has in hands such highly uncertain prior information m_0.

3.3 Impact of no prior information

Example 5: Let's examine the result of the fusion process if one doesn't include[5] the prior information $m_0(.)$ and if we combine directly only the three sources $m_1 \oplus m_2 \oplus m_3$ altogether with PCR5 or PCR6.

focal element	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	0	0.3	0
$\theta_6 \cup \theta_8$	0.75	0	0.25
$\overline{\theta_4 \cup \theta_8}$	0	0.7	0
I_t	0.25	0	0.75

Table 35: Quantitative inputs of VBIED problem.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_6	0.56875	0.56875
$\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$	0.13125	0.13125
θ_8	0.24375	0.24375
$\theta_4 \cup \theta_8$	0.05625	0.05625

Table 36: Result of $m_1 \oplus m_2 \oplus m_3$ for Table 35.

[5]Or equivalently we can take m_0 as the vacuous bba corresponding to $m_0(I_t) = 1$ and to the fully ignorant prior source.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0002	0.0002
θ_2	0.0002	0.0002
θ_3	0.0002	0.0002
θ_4	0.0002	0.0002
θ_5	0.0002	0.0002
θ_6	0.6989	0.6989
θ_7	0.0002	0.0002
θ_8	0.2998	0.2998

Table 37: $DSmP_\epsilon$ of $m_1 \oplus m_2 \oplus m_3$ for Table 35.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0219	0.0219
θ_2	0.0219	0.0219
θ_3	0.0219	0.0219
θ_4	0.0281	0.0281
θ_5	0.0219	0.0219
θ_6	0.5906	0.5906
θ_7	0.0219	0.0219
θ_8	0.2719	0.2719

Table 38: $BetP$ of $m_1 \oplus m_2 \oplus m_3$ for Table 35.

One gets $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.1875$ for $\theta_6 \cup \theta_7 \cup \theta_8$ with

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.8125, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.1875]$$

for $\theta_7 \cup \theta_8$, one has also $\Delta(\theta_7 \cup \theta_8) = 0.1875$ with

$$P(\theta_7 \cup \theta_8) \in [0.24375, 0.43125]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0.56875, 0.75625]$$

and for θ_8, one gets $\Delta(\theta_8) = 0.05625$ with

$$P(\theta_8) \in [0.24375, 0.3], \quad P(\bar{\theta}_8) \in [0.7, 0.75625]$$

The result presented in Table 36 is obviously the same as the one we obtain by combining the sources $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ altogether when taking $m_0(.)$ as the vacuous belief assignment, i.e. when $m_0(I_t) = 1$.

- **Answer to Q1:** Based on results of Tables 36-38 the decision based on max of belief, max of plausibility on either θ_8 or $\theta_7 \cup \theta_8$ is to NOT evacuate building B. Same conclusions is obtained when analyzing values of DSmP or BetP of θ_8 or $\theta_7 \cup \theta_8$. However, if we adopt the most prudent strategy based on decision-support hypothesis $\theta_6 \cup \theta_7 \cup \theta_8$ the decision will be to evacuate the building B since $DSmP(\theta_6 \cup \theta_7 \cup \theta_8) = 0.9989$. So we see the strong impact of the miss of prior information in the decision-making support process (by comparison between Example 1 and this example) when adopting more risky strategies for decision-making based either on $\theta_7 \cup \theta_8$ or on θ_8 only.

Let's compare now the source m_2 with respect to the $m_1 \oplus m_3$ fusion system when no prior information m_0 is used. Naturally, there is no need to fusion m_2 since we consider it alone. One has $\Delta_2(\theta_6 \cup \theta_7 \cup \theta_8) = \Delta_2(\theta_7 \cup \theta_8) = 1$ (i.e. the full imprecision on $P(\theta_6 \cup \theta_7 \cup \theta_8)$ and on $P(\theta_7 \cup \theta_8)$) whereas $\Delta_2(\theta_8) = 0.3$ with

$$P(\theta_8) \in [0, 0.3], \quad P(\bar{\theta}_8) \in [0.7, 1]$$

DSmP and BetP of $m_2(.)$ are the same since there is no singleton as focal element for $m_2(.)$ - see Table 39.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$BetP(.)$
θ_1	0.1167	0.1167
θ_2	0.1167	0.1167
θ_3	0.1167	0.1167
θ_4	0.1500	0.1500
θ_5	0.1167	0.1167
θ_6	0.1167	0.1167
θ_7	0.1167	0.1167
θ_8	0.1500	0.1500

Table 39: $DSmP_\epsilon$ and $BetP$ of m_2.

Based on max of Bel or max of Pl criteria on θ_8 (optimistic/risky strategy) the decision using m_2 (ANPR system alone) should be to NOT evacuate the building B. No decision can be taken using decision-support hypotheses $\theta_6 \cup \theta_7 \cup \theta_8$ or $\theta_7 \cup \theta_8$, nor on DSmP or BetP values since there is an ambiguity between θ_8 and θ_4.

Now if we combine m_1 with m_3 using PCR5 or PCR6 we get[6] results given in Tables 40 and 41.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
$\theta_6 \cup \theta_8$	0.8125	0.8125
I_t	0.1875	0.1875

Table 40: Result of $m_1 \oplus m_3$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0234	0.0234
θ_2	0.0234	0.0234
θ_3	0.0234	0.0234
θ_4	0.0234	0.0234
θ_5	0.0234	0.0234
θ_6	0.4297	0.4297
θ_7	0.0234	0.0234
θ_8	0.4297	0.4297

Table 41: $DSmP_\epsilon$ of $m_1 \oplus m_3$.

The values of $BetP(.)$ are same as those of $DSmP(.)$ because there is no singleton as focal element of $m_1 \oplus m_3$.

[6]Note that for two sources, PCR6 equals PCR5 [6], Vol.2.

Whence $\Delta_{13}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.1875$ with

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.8125, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.1875]$$

and $\Delta_{13}(\theta_7 \cup \theta_8) = 1$ with

$$P(\theta_7 \cup \theta_8) \in [0, 1], \qquad P(\overline{\theta_7 \cup \theta_8}) \in [0, 1]$$

and also $\Delta_{13}(\theta_8) = 1$ with

$$P(\theta_8) \in [0, 1], \qquad P(\bar{\theta}_8) \in [0, 1]$$

Based on max of Bel or max of Pl criteria on $\theta_6 \cup \theta_7 \cup \theta_8$, the decision using $m_1 \oplus m_3$ must be the evacuation of the building B. Same decision must be drawn when using DSmP results according to Table 41. No decision can be drawn based only on θ_8 or on $\theta_7 \cup \theta_8$ because of full imprecision on their corresponding probabilities.

- **Answer to Q2:** Similar remarks and conclusions to those given in Example 3 held also for Example 5, i.e. it is better to trust the most precise source for the most prudent decision-support hypothesis, and to decide to evacuate the building B if one has no prior information rather than using only information based on APNR system.

Example 6: It can be easily verified that the same analysis, remarks and conclusions for Q1 and Q2 as for Example 5 also hold when considering the sources m_1 and m_3 corresponding to the following input Table

focal element	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	0	0.3	0
$\theta_6 \cup \theta_8$	0.90	0	0.10
$\overline{\theta_4 \cup \theta_8}$	0	0.7	0
I_t	0.10	0	0.90

Table 42: Quantitative inputs of VBIED problem.

3.4 Impact of reliability of sources

The reliability of sources (when known) can be easily taken into using Shafer's classical discounting technique [5], p. 252, which consists in multiplying the masses of focal elements by the reliability factor α, and transferring all the remaining discounted mass to the full ignorance Θ. When $\alpha < 1$, such very simple reliability discounting technique discounts all focal elements with the same factor α and it increases the non specificity of the discounted sources since the mass committed to the full ignorance always increases. When $\alpha = 1$, no reliability discounting occurs (the bba is kept unchanged). Mathematically, Shafer's discounting technique for taking into account the reliability factor $\alpha \in [0, 1]$ of a

given source with a bba $m(.)$ and a frame Θ is defined by:

$$\begin{cases} m_\alpha(X) = \alpha \cdot m(X), & \text{for } X \neq \Theta \\ m_\alpha(\Theta) = \alpha \cdot m(\Theta) + (1-\alpha) \end{cases} \quad (1)$$

Example 7: Let's consider back the inputs of Table 28. The impact of strong unreliability of prior information m_0 has already been analyzed in Examples 3 and 4 by considering actually $\alpha_0 = 0.1$. Here we analyze the impact of reliabilities of sources m_0, m_1, m_2 and m_3 according presentation done in section 2.4 and we choose the following set of reliability factors $\alpha_0 = 0.9$, $\alpha_1 = 0.75$, $\alpha_2 = 0.75$ and $\alpha_3 = 0.25$. These values have been chosen approximatively but they reflect the fact that one has a very good confidence in our prior information, a good confidence in sources 1 and 2, and a low confidence in source 3. Let's examine the change in the fusion result of sources. Applying reliability discounting technique [5], the new inputs corresponding to the discounted bba's by (1) are given in Table 43.

focal element	$m_{\alpha_0}(.)$	$m_{\alpha_1}(.)$	$m_{\alpha_2}(.)$	$m_{\alpha_3}(.)$
$\theta_4 \cup \theta_8$	0.90	0	0.2250	0
$\theta_6 \cup \theta_8$	0	0.5625	0	0.0625
$\overline{\theta_4 \cup \theta_8}$	0	0	0.5250	0
I_t	0.10	0.4375	0.2500	0.9375

Table 43: Discounted inputs with $\alpha_0 = 0.9$, $\alpha_1 = 0.75$, $\alpha_2 = 0.75$ and $\alpha_3 = 0.25$.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_6	0.030967	0.030967
$\overline{\theta_4 \cup \theta_8}$	0.13119	0.11037
θ_8	0.26543	0.26543
$\theta_4 \cup \theta_8$	0.37256	0.33686
$\theta_6 \cup \theta_8$	0.063483	0.068147
I_t	0.13637	0.18822

Table 44: Result of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 43.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0040	0.0036
θ_2	0.0040	0.0036
θ_3	0.0040	0.0036
θ_4	0.0018	0.0019
θ_5	0.0040	0.0036
θ_6	0.1655	0.1535
θ_7	0.0040	0.0036
θ_8	0.8126	0.8266

Table 45: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 43.

From the fusion result of Table 44, one gets when considering $\theta_6 \cup \theta_7 \cup \theta_8$:

- with PCR5: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.64012$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.35988, 1]$$

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0389	0.0419
θ_2	0.0389	0.0419
θ_3	0.0389	0.0419
θ_4	0.2033	0.1920
θ_5	0.0389	0.0419
θ_6	0.1016	0.1070
θ_7	0.0389	0.0419
θ_8	0.5005	0.4915

Table 46: $BetP$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 43.

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.64012]$$

- with PCR6: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.635456$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.364544, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.635456]$$

and when considering $\theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_7 \cup \theta_8) = 0.703603$

$$P(\theta_7 \cup \theta_8) \in [0.26543, 0.969033]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0.030967, 0.73457]$$

- with PCR6: $\Delta(\theta_7 \cup \theta_8) = 0.703603$

$$P(\theta_7 \cup \theta_8) \in [0.26543, 0.969033]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0.030967, 0.73457]$$

and when considering θ_8 only

- with PCR5: $\Delta(\theta_8) = 0.572413$

$$P(\theta_8) \in [0.26543, 0.837843]$$

$$P(\bar{\theta}_8) \in [0.162157, 0.73457]$$

- with PCR6: $\Delta(\theta_8) = 0.593233$

$$P(\theta_8) \in [0.26543, 0.858663]$$

$$P(\bar{\theta}_8) \in [0.141337, 0.73457]$$

Using DSmP transformation, one gets high probabilities in $\theta_6 \cup \theta_7 \cup \theta_8$, $\theta_7 \cup \theta_8$ and in θ_8 because

$$DSmP_{\epsilon,PCR5}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.9821$$
$$DSmP_{\epsilon,PCR6}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.9837$$
$$DSmP_{\epsilon,PCR5}(\theta_7 \cup \theta_8) = 0.8166$$
$$DSmP_{\epsilon,PCR6}(\theta_7 \cup \theta_8) = 0.8302$$
$$DSmP_{\epsilon,PCR5}(\theta_8) = 0.8126$$
$$DSmP_{\epsilon,PCR6}(\theta_8) = 0.8266$$

- **Answer to Q1:** Based on these results, one sees that the decision to take is to evacuate the building B since one gets a high probability in decision-support hypotheses. Same conclusion is drawn when using max of Bel of max of Pl criteria. So there is a little impact of reliability discounting on the final decision with respect to Example 1. It is however worth to note that introducing reliability discounting increases the non specificity of information since now I_t is a new focal element of m_{α_0} and of m_{α_2} and in the final result we get the new focal element θ_6 appearing with PCR5 or PCR6 fusion rules. This θ_6 focal element doesn't exist in Example 1 when no reliability discounting is used. The decision to take in this case is to: **Evacuate the building** B.

To answer to the question Q2 for this Example 7, let's compute the fusion results of the fusion $m_0 \oplus m_2$ and $m_0 \oplus m_1 \oplus m_3$ using inputs given in Table 43. The fusion results with corresponding DSmP are given in the Tables 47-50.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
$\theta_4 \cup \theta_8$	0.74842	0.74842
$\overline{\theta_4 \cup \theta_8}$	0.22658	0.22658
I_t	0.025	0.025

Table 47: Result of $m_0 \oplus m_2$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0409	0.0409
θ_2	0.0409	0.0409
θ_3	0.0409	0.0409
θ_4	0.3773	0.3773
θ_5	0.0409	0.0409
θ_6	0.0409	0.0409
θ_7	0.0409	0.0409
θ_8	0.3773	0.3773

Table 48: $DSmP_\epsilon$ of $m_0 \oplus m_2$.

The result of BetP transformation is the same as with DSmP transformation since there is no singleton element as focal element of the resulting bba's when using PCR5 or PCR6 fusion rule.

Based on $m_0 \oplus m_2$ fusion result, one gets a large imprecision[7] on $P(\theta_8)$ since $\Delta_{02}(\theta_8) = 0.77342$ with

$$P(\theta_8) \in [0, 0.77342]$$

$$P(\bar{\theta}_8) \in [0.22658, 1]$$

[7] This imprecision is larger than in Example 1 which is normal because one has degraded the information of both prior and the source m_2.

and one gets a total imprecision when considering either $\theta_7 \cup \theta_8$ or $\theta_6 \cup \theta_7 \cup \theta_8$ since

$$P(\theta_7 \cup \theta_8) \in [0,1], \quad P(\overline{\theta_7 \cup \theta_8}) \in [0,1]$$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0,1], \quad P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0,1]$$

Based on max of bel or max of Pl criteria the decision using $m_0 \oplus m_2$ (i.e. with discounted sources m_0 and ANPR system m_2) should be to NOT evacuate the building B when working with decision-support hypothesis θ_8. No clear decision can be taken when working with $\theta_6 \cup \theta_7 \cup \theta_8$ or $\theta_7 \cup \theta_8$. Ambiguity in decision-making occurs between $\theta_8 = (A, V, B)$ and $\theta_4 = (A, V, \bar{B})$ when using DSmP or BetP transformations.

Let's examine now the result of the $m_0 \oplus m_1 \oplus m_3$ fusion given in Tables 49 and 50.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_8	0.53086	0.53086
$\theta_4 \cup \theta_8$	0.36914	0.36914
$\theta_6 \cup \theta_8$	0.058984	0.058984
I_t	0.041016	0.041016

Table 49: Result of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0001	0.0001
θ_2	0.0001	0.0001
θ_3	0.0001	0.0001
θ_4	0.0008	0.0008
θ_5	0.0001	0.0001
θ_6	0.0002	0.0002
θ_7	0.0001	0.0001
θ_8	0.9987	0.9987

Table 50: $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_3$.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0051	0.0051
θ_2	0.0051	0.0051
θ_3	0.0051	0.0051
θ_4	0.1897	0.1897
θ_5	0.0051	0.0051
θ_6	0.0346	0.0346
θ_7	0.0051	0.0051
θ_8	0.7500	0.7500

Table 51: $BetP$ of $m_0 \oplus m_1 \oplus m_3$.

One sees clearly the impact of reliability discounting on the specificity of information provided by the fusion of sources. Indeed when using the discounting of sources (mainly because we introduce I_t as focal element for m_0) one gets now 4 focal elements whereas we did get only two focal elements when no discounting

was used (see Table 9). Based on $m_0 \oplus m_1 \oplus m_3$ fusion result, one gets $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.410156$ with

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.589844, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.410156]$$

and also $\Delta_{013}(\theta_7 \cup \theta_8) = 0.46914$ with

$$P(\theta_7 \cup \theta_8) \in [0.53086, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.46914]$$

and $\Delta_{013}(\theta_8) = 0.46914$ with

$$P(\theta_8) \in [0.53086, 1]$$

$$P(\bar{\theta}_8) \in [0, 0.46914]$$

Based on max of Bel or max of Pl, the decision using $m_0 \oplus m_1 \oplus m_3$ should be to evacuate the building B. Same decision would be taken based on DSmP values. It is worth to note that the precision on the result obtained with $m_0 \oplus m_1 \oplus m_3$ is much better than with $m_0 \oplus m_2$ since $\Delta_{013}(\theta_8) < \Delta_{02}(\theta_8)$, or $\Delta_{013}(\theta_7 \cup \theta_8) < \Delta_{02}(\theta_7 \cup \theta_8)$, or $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) < \Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8)$. Moreover it is easy to verify that $m_0 \oplus m_1 \oplus m_3$ fusion system is more informative than $m_0 \oplus m_2$ fusion system because Shannon entropy of DSmP of $m_0 \oplus m_2$ is much bigger than Shannon entropy of DSmP of $m_0 \oplus m_1 \oplus m_3$.

- **Answer to Q2:** Since the information obtained by the fusion $m_0 \oplus m_2$ is less informative and less precise than the information obtained with the fusion $m_0 \oplus m_1 \oplus m_3$, it is better to choose and to trust the fusion system $m_0 \oplus m_1 \oplus m_3$ rather than $m_0 \oplus m_2$. Based on this choice, the final decision will be to evacuate the building B.

3.5 Impact of importance of sources

The importance discounting technique has been proposed recently by the authors in [7] and consists in discounting the masses of focal elements by a factor $\beta \in [0, 1]$ and in transferring the remaining mass to empty set, i.e.

$$\begin{cases} m_\beta(X) = \beta \cdot m(X), & \text{for } X \neq \emptyset \\ m_\beta(\emptyset) = \beta \cdot m(\emptyset) + (1 - \beta) \end{cases} \quad (2)$$

It has been proved in [7] that such importance discounting technique preserves the specificity of the information and that Dempster's rule of combination doesn't respond to such new interesting discounting technique specially useful and crucial in multicriteria decision-making support.

In the extreme case, the method proposed in [7] reinforces the highest mass of the focal element of the source having the biggest importance factor as soon as the other sources have their importance factors tending towards zero. This reinforcement may be a disputable behavior. To avoid such behavior, we propose here to use the same importance discounting technique, but the fusion of discounted sources is done a bit differently in three steps as follows:

- Step 1: Discount each source with its importance discounting factor according to (2).
- Step 2: Apply PCR5 or PCR6 fusion rule with unnormalized discounted bba's, i.e. as if the discounted mass committed to empty set for each source was zero.
- Step 3: Normalize the result to get the sum of masses of focal elements to be one.

Let's examine the impact of the importance of the sources in the fusion process for final decision-making through the next very simple illustrating example.

Example 8: To evaluate this we consider the same inputs as in Table 1 and we consider that source 1 (Analyst 1 with 10 years experience) is much more important than source 3 (Analyst 2 with no experience). To reflect the difference between importance of this sources we consider the following relative importance factors $\beta_1 = 0.9$ and $\beta_3 = 0.5$. We also assume that source 0 (prior information) and source 2 (ANPR system) have the same maximal importance, i.e. $\beta_0 = \beta_2 = 1$, i.e $\beta = (\beta_0, \beta_1, \beta_2, \beta_3) = (1, 0.9, 1, 0.5)$. These values have been chosen approximatively but they do reflect the fact that sources m_0 and m_2 have same importance in the fusion process, and that sources m_1 and m_3 may have less importance in the fusion process taking into the fact that m_3 is considered as less important than m_1. Let's examine the change in the fusion result of sources in this example with respect to what we get in Example 1.

In applying importance discounting technique [7] with the aforementioned fusion approach, the new inputs corresponding to the discounted (unnormalized) bba's by (2) are given in Table 52.

focal element	$m_{\beta_0}(.)$	$m_{\beta_1}(.)$	$m_{\beta_2}(.)$	$m_{\beta_3}(.)$
$\emptyset$	0	0.1	0	0.5
$\theta_4 \cup \theta_8$	1	0	0.70	0
$\theta_6 \cup \theta_8$	0	0.675	0	0.125
$\overline{\theta_4 \cup \theta_8}$	0	0	0.30	0
I_t	0	0.225	0	0.375

Table 52: Discounted inputs with $\beta_0 = 1$, $\beta_1 = 0.9$, $\beta_2 = 1$ and $\beta_3 = 0.5$.

and the fusion result is given in Table 53.

As we can see, the importance discounting doesn't degrade the specificity of sources since no mass is

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_8	0.24375	0.24375
$\theta_4 \cup \theta_8$	0.36788	0.33034
$\theta_6 \cup \theta_8$	0.10552	0.14132
$\overline{\theta_4 \cup \theta_8}$	0.21814	0.19186
I_t	0.064701	0.092734

Table 53: Result of $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_2} \oplus m_{\beta_3}$.

committed to partial ignorances, and it doesn't also increase the number of focal elements of the resulting bba contrariwise to the reliability discounting approach. Indeed in Table 53 one gets only 5 focal elements whereas one gets 6 focal elements with reliability discounting as shown in Table 44 of Example 7. The corresponding DSmP and BetP values of bba's given in Table 53 are summarized in Tables 54 and 55 .

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0.0366	0.0323
θ_2	0.0366	0.0323
θ_3	0.0366	0.0323
θ_4	0.0018	0.0017
θ_5	0.0366	0.0323
θ_6	0.0370	0.0329
θ_7	0.0366	0.0323
θ_8	0.7781	0.8036

Table 54: $DSmP_\epsilon$ of $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_2} \oplus m_{\beta_3}$.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0.0444	0.0436
θ_2	0.0444	0.0436
θ_3	0.0444	0.0436
θ_4	0.1920	0.1768
θ_5	0.0444	0.0436
θ_6	0.0972	0.1142
θ_7	0.0444	0.0436
θ_8	0.4885	0.4912

Table 55: $BetP$ of $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_2} \oplus m_{\beta_3}$.

From the fusion result of Table 53, one gets when considering $\theta_6 \cup \theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.65073$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.34927, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.65073]$$

- with PCR6: $\Delta(\theta_6 \cup \theta_7 \cup \theta_8) = 0.61493$

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.38507, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.61493]$$

and when considering $\theta_7 \cup \theta_8$

- with PCR5: $\Delta(\theta_7 \cup \theta_8) = 0.75625$

$$P(\theta_7 \cup \theta_8) \in [0.24375, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.75625]$$

- with PCR6: $\Delta(\theta_7 \cup \theta_8) = 0.75625$

$$P(\theta_7 \cup \theta_8) \in [0.24375, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.75625]$$

and when considering θ_8 only

- with PCR5: $\Delta(\theta_8) = 0.53811$

$$P(\theta_8) \in [0.24375, 0.78186]$$

$$P(\bar{\theta}_8) \in [0.21814, 0.75625]$$

- with PCR6: $\Delta(\theta_8) = 0.564393$

$$P(\theta_8) \in [0.24375, 0.80814]$$

$$P(\bar{\theta}_8) \in [0.19186, 0.75625]$$

Using DSmP transformation, one gets a high probability in decision-support hypotheses because

$$DSmP_{\epsilon,PCR5}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.8517$$
$$DSmP_{\epsilon,PCR6}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.8688$$
$$DSmP_{\epsilon,PCR5}(\theta_7 \cup \theta_8) = 0.8147$$
$$DSmP_{\epsilon,PCR6}(\theta_7 \cup \theta_8) = 0.8359$$
$$DSmP_{\epsilon,PCR5}(\theta_8) = 0.7781$$
$$DSmP_{\epsilon,PCR6}(\theta_8) = 0.8036$$

- **Answer to Q1:** Based on these results, one sees that the decision to take based either on max of Bel or max of Pl on $\theta_6 \cup \theta_7 \cup \theta_8$, on $\theta_7 \cup \theta_8$ or on θ_8 only, or also based on DSmP, is to evacuate the building B.

To answer to the question Q2 for this Example 8, let's compute the fusion results of the fusion $m_{\beta_0} \oplus m_{\beta_2}$ and $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$ using inputs given in Table 52. Because one has considered $\beta_0 = \beta_2 = 1$, one does not actually discount sources m_0 and m_2 and therefore the $m_0 \oplus m_2$ fusion results are already given in Tables 5 and 6 of Example 1. Therefore based on max of Bel or max of Pl criteria on θ_8 the decision using $m_0 \oplus m_2$ is to NOT evacuate the building B since $P(\theta_8) \in [0, 0.71176]$ and $P(\bar{\theta}_8) \in [0.28824, 1]$ and $\Delta_{02}(\theta_8) = 0.71176$. Same decision would be taken based on DSmP values with the $m_0 \oplus m_2$ fusion sub-system.

Let's now compute the fusion $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$ with the importance discounted sources $m_{\beta_0=1} = m_0$, m_{β_1} and m_{β_3}. The fusion results are given in Tables 56, 57 and 58.

focal element	$m_{PCR5}(.)$	$m_{PCR6}(.)$
θ_8	0.8125	0.8125
$\theta_4 \cup \theta_8$	0.1875	0.1875

Table 56: Result of $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$.

Singletons	$DSmP_{\epsilon,PCR5}(.)$	$DSmP_{\epsilon,PCR6}(.)$
θ_1	0	0
θ_2	0	0
θ_3	0	0
θ_4	0.0002	0.0002
θ_5	0	0
θ_6	0	0
θ_7	0	0
θ_8	0.9998	0.9998

Table 57: $DSmP_\epsilon$ of $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$.

Singletons	$BetP_{PCR5}(.)$	$BetP_{PCR6}(.)$
θ_1	0	0
θ_2	0	0
θ_3	0	0
θ_4	0.0002	0.0002
θ_5	0	0
θ_6	0	0
θ_7	0	0
θ_8	0.9998	0.9998

Table 58: $BetP$ of $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$.

Based on $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$ fusion result, one gets same result with PCR5 and PCR6 in this case, and one gets $\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.1875$ with

$$P(\theta_6 \cup \theta_7 \cup \theta_8) \in [0.8125, 1]$$

$$P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) \in [0, 0.1875]$$

and $\Delta_{013}(\theta_7 \cup \theta_8) = 0.1875$ with

$$P(\theta_7 \cup \theta_8) \in [0.8125, 1]$$

$$P(\overline{\theta_7 \cup \theta_8}) \in [0, 0.1875]$$

and also $\Delta_{013}(\theta_8) = 0.1875$ with

$$P(\theta_8) \in [0.8125, 1]$$

$$P(\bar{\theta}_8) \in [0, 0.1875]$$

Based on max of Bel or max of Pl, the decision taken using $\theta_6 \cup \theta_7 \cup \theta_8$, $\theta_7 \cup \theta_8$ or θ_8 for the $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$ fusion sub-system should be to evacuate the building B. Same decision must be taken based on DSmP or BetP values. It is worth to note that the precision[8] on the result obtained with subsystem $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$ is much better than with subsystem $m_{\beta_0=1} \oplus m_{\beta_2=1}$ since $(\Delta_{013}(\theta_8) = 0.1875) < (\Delta_{02}(\theta_8) = 0.71176)$, and also $(\Delta_{013}(\theta_7 \cup \theta_8) = 0.1875) < (\Delta_{02}(\theta_7 \cup \theta_8) = 1)$, and $(\Delta_{013}(\theta_6 \cup \theta_7 \cup \theta_8) = 0.1875) < (\Delta_{02}(\theta_6 \cup \theta_7 \cup \theta_8) = 1)$.

[8] see Example 1 for the numerical results of $m_{\beta_0=1} \oplus m_{\beta_2=1}$.

- **Answer to Q2:** The analysis of both fusion subsystems $m_{\beta_0=1} \oplus m_{\beta_2=1}$ and $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$ shows that the $m_{\beta_0} \oplus m_{\beta_1} \oplus m_{\beta_3}$ subsystem must be chosen because it provides the most precise results and therefor the decision will be to evacuate the building B whatever the decision-support hypothesis is chosen $\theta_6 \cup \theta_7 \cup \theta_8$, $\theta_7 \cup \theta_8$ or θ_8.

3.6 Using imprecise bba's

Let's examine the fusion result when dealing directly with imprecise bba's. We just consider here a simple imprecise example which considers both inputs of Examples 1 and 2 to generate imprecise bba's inputs.

Example 9: We consider the imprecise bba's according to input Table 59.

focal element	$m_0(.)$	$m_1(.)$	$m_2(.)$	$m_3(.)$
$f_1 = \theta_4 \cup \theta_8$	1	0	0.3	0
$f_2 = \theta_6 \cup \theta_8$	0	[0.75,0.9]	0	[0.10,0.25]
$f_3 = \overline{\theta_4 \cup \theta_8}$	0	0	0.7	0
I_t	0	[0.1,0.25]	0	[0.75,0.9]

Table 59: Imprecise quantitative inputs for VBIED problem.

Applying the conjunctive rule, we have $1 \times 2 \times 2 \times 2 = 8$ products to compute which are listed below:

- Product $\pi_1 = 1 \boxdot [0.75, 0.90] \boxdot 0.3 \boxdot [0.10, 0.25]$. Using operators on sets defined in [6], Vol.1, Chap. 6, one gets $\pi_1 = [0.75, 0.90] \boxdot [0.03, 0.075] = [0.0225, 0.0675]$ which is committed to $f_1 \cap f_2 = \theta_8$.
- Product $\pi_2 = 1 \boxdot [0.75, 0.90] \boxdot 0.3 \boxdot [0.75, 0.90]$ is equal to $[0.75, 0.90] \boxdot [0.225, 0.27] = [0.16875, 0.243]$ which is also committed to $f_1 \cap f_2 = \theta_8$.
- Product $\pi_3 = 1 \boxdot [0.75, 0.90] \boxdot 0.7 \boxdot [0.10, 0.25] = [0.0525, 0.1575]$ corresponds to the imprecise mass of $f_1 \cap f_2 \cap f_3 = \emptyset$ which will be redistributed back to f_1, f_2 and f_3 according to PCR6.
- Product $\pi_4 = 1 \boxdot [0.75, 0.90] \boxdot 0.7 \boxdot [0.75, 0.90] = [0.39375, 0.567]$ corresponds to the imprecise mass of $f_1 \cap f_2 \cap f_3 \cap I_t = \emptyset$ which will be redistributed back to f_1, f_2, f_3 and I_t according to PCR6.
- Product $\pi_5 = 1 \boxdot [0.10, 0.25] \boxdot 0.3 \boxdot [0.10, 0.25] = [0.003, 0.01875]$ is committed to $f_1 \cap f_2 = \theta_8$.
- Product $\pi_6 = 1 \boxdot [0.10, 0.25] \boxdot 0.3 \boxdot [0.75, 0.90] = [0.0225, 0.0675]$ is committed to f_1.
- Product $\pi_7 = 1 \boxdot [0.10, 0.25] \boxdot 0.7 \boxdot [0.10, 0.25] = [0.007, 0.04375]$ corresponds to the imprecise mass of $f_1 \cap I_t \cap f_3 \cap f_2 = \emptyset$ which will be redistributed back to f_1, f_2, f_3 and I_t according to PCR6.

- Product $\pi_8 = 1 \boxdot [0.10, 0.25] \boxdot 0.7 \boxdot [0.75, 0.90] = [0.0525, 0.1575]$ corresponds to the imprecise mass of $f_1 \cap I_t \cap f_3 \cap I_t = \emptyset$ which will be redistributed back to f_1, f_3 and I_t according to PCR6.

We now redistribute the imprecise masses π_3, π_4, π_7 and π_8 associated with the empty set using PCR6 principle. Lets' compute the proportions of π_3, π_4, π_7 and π_8 committed to each focal element involved in the conflict they are associated with.

- The product $\pi_3 = [0.0525, 0.1575]$ is distributed to f_1, f_2 and f_3 according to PCR6 as follows

$$\begin{aligned}\frac{x_{f_1,\pi_3}}{1} &= \frac{y_{f_2,\pi_3}}{[0.75, 0.90] \boxplus [0.10, 0.25]} = \frac{z_{f_3,\pi_3}}{0.7} \\ &= \frac{\pi_3}{1 \boxplus [0.75, 0.90] \boxplus [0.10, 0.25] \boxplus 0.7} \\ &= \frac{[0.0525, 0.1575]}{1.7 \boxplus [0.85, 1.15]} = \frac{[0.0525, 0.1575]}{[2.55, 2.85]} \\ &= [\frac{0.0525}{2.85}, \frac{0.1575}{2.55}] \\ &= [0.018421, 0.061765]\end{aligned}$$

whence

$$\begin{aligned}x_{f_1,\pi_3} &= 1 \boxdot [0.018421, 0.061765] \\ &= [0.018421, 0.061765] \\ y_{f_2,\pi_3} &= [0.85, 1.15] \boxdot [0.018421, 0.061765] \\ &= [0.015658, 0.071029] \\ z_{f_3,\pi_3} &= 0.7 \boxdot [0.018421, 0.061765] \\ &= [0.012895, 0.043236]\end{aligned}$$

- The product $\pi_4 = [0.39375, 0.567]$ is distributed to f_1, f_2, f_3 and I_t according to PCR6 as follows

$$\begin{aligned}\frac{x_{f_1,\pi_4}}{1} &= \frac{y_{f_2,\pi_4}}{[0.75, 0.90]} = \frac{z_{f_3,\pi_4}}{0.7} = \frac{w_{I_t,\pi_4}}{[0.75, 0.90]} \\ &= \frac{\pi_4}{1 \boxplus [0.75, 0.90] \boxplus 0.7 \boxplus [0.75, 0.90]} \\ &= \frac{[0.39375, 0.567]}{[3.2, 3.5]} \\ &= [0.1125, 0.177188]\end{aligned}$$

whence

$$\begin{aligned}x_{f_1,\pi_4} &= 1 \boxdot [0.1125, 0.177188] \\ &= [0.1125, 0.177188] \\ y_{f_2,\pi_4} &= [0.75, 0.90] \boxdot [0.1125, 0.177188] \\ &= [0.084375, 0.159469] \\ z_{f_3,\pi_4} &= 0.7 \boxdot [0.1125, 0.177188] \\ &= [0.07875, 0.124031] \\ w_{I_t,\pi_4} &= [0.75, 0.90] \boxdot [0.1125, 0.177188] \\ &= [0.084375, 0.159469]\end{aligned}$$

- The product $\pi_7 = [0.007, 0.04375]$ is distributed to f_1, f_2, f_3 and I_t according to PCR6 as follows

$$\begin{aligned}\frac{x_{f_1,\pi_7}}{1} &= \frac{y_{f_2,\pi_7}}{[0.10, 0.25]} = \frac{z_{f_3,\pi_7}}{0.7} = \frac{w_{I_t,\pi_7}}{[0.10, 0.25]} \\ &= \frac{\pi_7}{1 \boxplus [0.10, 0.25] \boxplus 0.7 \boxplus [0.10, 0.25]} \\ &= \frac{[0.007, 0.04375]}{[1.9, 2.2]} \\ &= [0.003182, 0.023026]\end{aligned}$$

whence

$$\begin{aligned}x_{f_1,\pi_7} &= 1 \boxdot [0.003182, 0.023026] \\ &= [0.003182, 0.023026] \\ y_{f_2,\pi_7} &= [0.10, 0.25] \boxdot [0.003182, 0.023026] \\ &= [0.000318, 0.005757] \\ z_{f_3,\pi_7} &= 0.7 \boxdot [0.003182, 0.023026] \\ &= [0.002227, 0.016118] \\ w_{I_t,\pi_7} &= [0.10, 0.25] \boxdot [0.003182, 0.023026] \\ &= [0.000318, 0.005757]\end{aligned}$$

- The product $\pi_8 = [0.0525, 0.1575]$ is distributed to f_1, f_3 and I_t according to PCR6 as follows

$$\begin{aligned}\frac{x_{f_1,\pi_8}}{1} &= \frac{z_{f_3,\pi_8}}{0.7} = \frac{w_{I_t,\pi_8}}{[0.10, 0.25] \boxplus [0.75, 0.90]} \\ &= \frac{\pi_8}{1 \boxplus 0.7 \boxplus [0.10, 0.25] \boxplus [0.75, 0.90]} \\ &= \frac{[0.0525, 0.1575]}{[2.55, 2.85]} \\ &= [0.018421, 0.061765]\end{aligned}$$

whence

$$\begin{aligned}x_{f_1,\pi_8} &= 1 \boxdot [0.018421, 0.061765] \\ &= [0.018421, 0.061765] \\ z_{f_3,\pi_8} &= 0.7 \boxdot [0.018421, 0.061765] \\ &= [0.012895, 0.043235] \\ w_{I_t,\pi_8} &= [0.85, 1.15] \boxdot [0.018421, 0.061765] \\ &= [0.015658, 0.071029]\end{aligned}$$

Summing the results, we get for $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ with PCR6 the following imprecise m_{PCR6} bba:

$$\begin{aligned}m_{PCR6}(\theta_8) &= \pi_1 \boxplus \pi_2 \boxplus \pi_5 \\ &= [0.19425, 0.32925] \\ m_{PCR6}(f_1) &= x_{f_1,\pi_3} \boxplus x_{f_1,\pi_4} \boxplus x_{f_1,\pi_7} \boxplus x_{f_1,\pi_8} \\ &= [0.152524, 0.323743] \\ m_{PCR6}(f_2) &= y_{f_2,\pi_3} \boxplus y_{f_2,\pi_4} \boxplus y_{f_2,\pi_7} \\ &= [0.100351, 0.236255] \\ m_{PCR6}(f_3) &= z_{f_3,\pi_3} \boxplus z_{f_3,\pi_4} \boxplus z_{f_3,\pi_7} \boxplus z_{f_3,\pi_8} \\ &= [0.106767, 0.226620] \\ m_{PCR6}(I_t) &= w_{I_t,\pi_4} \boxplus w_{I_t,\pi_7} \boxplus w_{I_t,\pi_8} \\ &= [0.100351, 0.236255]\end{aligned}$$

where $\boxplus$, $\boxdot$ and $\boxslash$ operators (i.e. the addition, multiplication and division of imprecise values), and other operators on sets, were defined in [6], Vol. 1, p127–130 by

$$S_1 \boxplus S_2 = \{x | x = s_1 + s_2, s_1 \in S_1, s_2 \in S_2\}$$

$$S_1 \boxdot S_2 = \{x | x = s_1 \cdot s_2, s_1 \in S_1, s_2 \in S_2\}$$

$$S_1 \boxslash S_2 = \{x | x = s_1 / s_2, s_1 \in S_1, s_2 \in S_2\}$$

with

$$\inf(S_1 \boxplus S_2) = \inf(S_1) + \inf(S_2)$$
$$\sup(S_1 \boxplus S_2) = \sup(S_1) + \sup(S_2)$$

and

$$\inf(S_1 \boxdot S_2) = \inf(S_1) \cdot \inf(S_2)$$
$$\sup(S_1 \boxdot S_2) = \sup(S_1) \cdot \sup(S_2)$$

and

$$\inf(S_1 \boxslash S_2) = \inf(S_1) / \sup(S_2)$$
$$\sup(S_1 \boxslash S_2) = \sup(S_1) / \inf \sup(S_2)$$

We have summarized the results in Table 60. The left column of this table corresponds to the imprecise values of m_{PCR6} based on exact calculus with operators on sets (i.e. the exact calculus with imprecision). The right column of this table (m_{PCR6}^{approx}) corresponds to the result obtained with non exact calculus based on results given in Examples 1 and 2 in right columns of Tables 2 and 12. This is what we call approximate results since they are not based on exact calculus with operators on sets. One shows an important differences between results in left and right columns which can make an impact on final decision process when working with imprecise bba's and we suggest to always use exact calculus (more complicated) instead of approximate calculus (more easier) in order to get the real imprecision on bba's values. Same approach can be done for combining imprecise bba's with PCR5 (not reported in this paper).

focal element	$m_{PCR6}(.)$	$m_{PCR6}^{approx}(.)$
θ_8	[0.194250,0.329250]	[0.24375,0.27300]
$f_1 = \theta_4 \cup \theta_8$	[0.152524,0.323743]	[0.23935,0.29641]
$f_2 = \theta_6 \cup \theta_8$	[0.100351,0.236255]	[0.14587, 0.16950]
$f_3 = \overline{\theta_4 \cup \theta_8}$	[0.106767,0.226620]	[0.14865,0.16811]
I_t	[0.100351,0.236255]	[0.14587, 0.16950]

Table 60: Results of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 59.

Based on results on left column of Table 60, one can easily compute the imprecise Bel and Pl values also of decision-support hypotheses which are

$$\begin{aligned} Bel(\theta_6 \cup \theta_7 \cup \theta_8) &= [0.294601, 0.565505] \\ Pl(\theta_6 \cup \theta_7 \cup \theta_8) &= [0.654243, 1.352123] \equiv [0.654243, 1] \\ Bel(\theta_7 \cup \theta_8) &= [0.194250, 0.329250] \\ Pl(\theta_7 \cup \theta_8) &= [0.654243, 1.352123] \equiv [0.654243, 1] \\ Bel(\theta_8) &= [0.194250, 0.329250] \\ Pl(\theta_8) &= [0.547476, 1.125502] \equiv [0.547476, 1] \end{aligned}$$

Therefore, one gets the following imprecision ranges for probabilities

$$\begin{aligned} P(\theta_6 \cup \theta_7 \cup \theta_8) &\in [0.294601, 1] \\ P(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) &\in [0, 0.705399] \\ P(\theta_7 \cup \theta_8) &\in [0.194250, 1] \\ P(\overline{\theta_7 \cup \theta_8}) &\in [0, 0.805750] \\ P(\theta_8) &\in [0.194250, 1] \\ P(\bar{\theta}_8) &\in [0, 0.805750] \end{aligned}$$

Based on max of Bel or max of Pl criteria, one sees that the decision will be to evacuate the building B whatever the decision-support hypothesis we prefer $\theta_6 \cup \theta_7 \cup \theta_8$, $\theta_7 \cup \theta_8$ or θ_8.

Let's compute now the imprecise DSmP values for $\epsilon = 0.001$. The focal element $f_1 = \theta_4 \cup \theta_8$ is redistributed back to θ_4 and θ_8 directly proportionally to their corresponding masses and cardinalities

$$\begin{aligned} \frac{x_{\theta_4}}{0 \boxplus 0.001} &= \frac{y'_{\theta_8}}{[0.194250, 0.329250] \boxplus 0.001} \\ &= \frac{m_{PCR6}(\theta_4 \cup \theta_8)}{0.002 \boxplus [0.194250, 0.329250]} \\ &= \frac{[0.152524, 0.323743]}{[0.196250, 0.331250]} \\ &= [\frac{0.152524}{0.331250}, \frac{0.323743}{0.196250}] \\ &= [0.46045, 1.64965] \end{aligned}$$

whence

$$\begin{aligned} x_{\theta_4} &= 0.001 \boxdot [0.46045, 1.64965] \\ &= [0.000460, 0.001650] \\ y'_{\theta_8} &= [0.195250, 0.330250] \boxdot [0.46045, 1.64965] \\ &= [0.089903, 0.544797] \end{aligned}$$

The focal element $f_2 = \theta_6 \cup \theta_8$ is redistributed back to θ_6 and θ_8 directly proportionally to their corresponding

masses and cardinalities

$$\begin{aligned}\frac{z_{\theta_6}}{0 \boxplus 0.001} &= \frac{y''_{\theta_8}}{[0.194250, 0.329250] \boxplus 0.001} \\ &= \frac{m_{PCR6}(\theta_6 \cup \theta_8)}{0.002 \boxplus [0.194250, 0.329250]} \\ &= \frac{[0.100351, 0.236255]}{[0.196250, 0.331250]} \\ &= [0.302943, 1.20385]\end{aligned}$$

whence

$$\begin{aligned}z_{\theta_6} &= 0.001 \boxdot [0.302943, 1.20385] \\ &= [0.000303, 0.001204] \\ y''_{\theta_8} &= [0.195250, 0.330250] \boxdot [0.302943, 1.20385] \\ &= [0.0591496, 0.3975714]\end{aligned}$$

The focal element $f_3 = \overline{\theta_4 \cup \theta_8}$ which is also equal to $\theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_5 \cup \theta_6 \cup \theta_7$ is redistributed back to θ_1, θ_2, θ_3, θ_5, θ_6 and θ_7 directly proportionally to their corresponding masses and cardinalities

$$\begin{aligned}\frac{w_{\theta_1}}{0 \boxplus 0.001} &= \frac{w_{\theta_2}}{0 \boxplus 0.001} = \frac{w_{\theta_3}}{0 \boxplus 0.001} = \frac{w_{\theta_5}}{0 \boxplus 0.001} \\ &= \frac{w_{\theta_6}}{0 \boxplus 0.001} = \frac{w_{\theta_7}}{0 \boxplus 0.001} \\ &= \frac{m_{PCR6}(\overline{\theta_4 \cup \theta_8})}{0.006} \\ &= \frac{[0.106767, 0.226620]}{0.006} \\ &= [17.7945, 37.770]\end{aligned}$$

Since all are equal, we get

$$\begin{aligned}w_{\theta_1} &= w_{\theta_2} = w_{\theta_3} = w_{\theta_5} = w_{\theta_6} = w_{\theta_7} \\ &= 0.001 \boxdot [17.7945, 37.770] \\ &= [0.0177945, 0.03777]\end{aligned}$$

The total ignorance $I_t = \theta_1 \cup \theta_2 \cup \theta_3 \cup \theta_4 \cup \theta_5 \cup \theta_6 \cup \theta_7 \cup \theta_8$ is redistributed back to all eight elements of the frame Θ directly proportionally to their corresponding masses and cardinalities

$$\begin{aligned}\frac{v_{\theta_1}}{0 \boxplus 0.001} &= \frac{v_{\theta_2}}{0 \boxplus 0.001} = \frac{v_{\theta_3}}{0 \boxplus 0.001} = \frac{v_{\theta_4}}{0 \boxplus 0.001} \\ &= \frac{v_{\theta_5}}{0 \boxplus 0.001} = \frac{v_{\theta_6}}{0 \boxplus 0.001} = \frac{v_{\theta_7}}{0 \boxplus 0.001} \\ &= \frac{v_{\theta_8}}{[0.194250, 0.329250] \boxplus 0.001} \\ &= \frac{m_{PCR6}(I_t)}{[0.194250, 0.329250] \boxplus 0.008} \\ &= \frac{[0.100351, 0.236255]}{[0.202550, 0.337250} \\ &= [0.297557, 1.16813]\end{aligned}$$

whence

$$\begin{aligned}v_{\theta_1} &= v_{\theta_2} = v_{\theta_3} = v_{\theta_4} = v_{\theta_5} = v_{\theta_6} = v_{\theta_7} \\ &= 0.001 \boxdot [0.297557, 1.16813] \\ &= [0.000298, 0.001168] \\ v_{\theta_8} &= [0.195250, 0.330250] \boxdot [0.297557, 1.16813] \\ &= [0.058098, 0.385775]\end{aligned}$$

The imprecise DSmP probabilities are computed by

$$\begin{aligned}DSmP_{\epsilon,PCR6}(\theta_1) &= w_{\theta_1} \boxplus v_{\theta_1} \\ DSmP_{\epsilon,PCR6}(\theta_2) &= w_{\theta_2} \boxplus v_{\theta_2} \\ DSmP_{\epsilon,PCR6}(\theta_3) &= w_{\theta_3} \boxplus v_{\theta_3} \\ DSmP_{\epsilon,PCR6}(\theta_4) &= x_{\theta_4} \boxplus v_{\theta_4} \\ DSmP_{\epsilon,PCR6}(\theta_5) &= w_{\theta_5} \boxplus v_{\theta_5} \\ DSmP_{\epsilon,PCR6}(\theta_6) &= z_{\theta_6} \boxplus w_{\theta_6} \boxplus v_{\theta_6} \\ DSmP_{\epsilon,PCR6}(\theta_7) &= w_{\theta_7} \boxplus v_{\theta_7} \\ DSmP_{\epsilon,PCR6}(\theta_8) &= y'_{\theta_8} \boxplus y''_{\theta_8} \boxplus v_{\theta_8}\end{aligned}$$

which are summarized[9] in Table 61 below.

Singletons	$DSmP_{\epsilon,PCR6}(.)$
θ_1	[0.0181,0.0389]
θ_2	[0.0181,0.0389]
θ_3	[0.0181,0.0389]
θ_4	[0.0008,0.0028]
θ_5	[0.0181,0.0389]
θ_6	[0.0184 0.0402]
θ_7	[0.0181,0.0389]
θ_8	[0.2072,1]

Table 61: Imprecise $DSmP_\epsilon$ of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 59.

- **Answer to Q1:** As we have shown, it is possible to fuse imprecise bba's with PCR6, and PCR5 too (see [6], Vol. 2) to get an imprecise result for decision-making support under uncertainty and imprecision. It is also possible to compute the exact imprecise values of DSmP if necessary. According to our analysis and our results, and using either the max of Bel, the max of Pl of the max of DSmP criterion, the decision will be to evacuate the building B. Of course, a similar analysis can be done to answer to the question Q2 when working with imprecise bba's, and for for computing imprecise BetP values as well.

4 Qualitative approach

In this section we just show how the fusion and decision-making can be done using qualitative information expressed with labels. In our previous examples the quantitative baa's have been defined ad-hoc in satisfying

[9] Actually for θ_8, one gets with exact calculus of imprecision $DSmP_{\epsilon,PCR6}(\theta_8) = [0.2072, 1.3281]$, but since a probability cannot be greater than 1, the upper bound of imprecision interval has been set to 1.

some reasonable modeling and using minimal assumption compatible with what is given in the statement of the VBIED problem. The numerical values can be slightly changed (as we have shown in Examples 1 and 2, or in Examples 3 or 4) or they can even be taken as imprecise as in Example 9, but they still need to be kept coherent with sources reports in order to obtain what we consider as pertinent and motivated and fully justified answers to questions Q1 and Q2.

In this section we show how to solve the problem using qualitative information using labels. We investigate the possibility to work either with a minimal set of labels $\{L_1 = Low, L_2 = High\}$ (i.e. with $m = 2$ labels) or a more refined set consisting in 3 labels $\{L_1 = Low, L_2 = Medium, L_3 = High\}$ (i.e. with $m = 3$ labels). Each set is extended with minimal L_0 and maximal L_{m+1} labels as follows (see [6], Vol.3, Chap. 2 for examples and details)

$$\mathcal{L}_2 = \{L_0 \equiv 0, L_1 = Low, L_2 = High, L_3 \equiv 1\}$$

and

$$\mathcal{L}_3 = \{L_0 \equiv 0, L_1 = Low, L_2 = Medium, L_3 = High, L_4 \equiv 1\}$$

To simplify the presentation, we only present the results when combining directly the sources altogether and considering that they have all the same maximal reliability and importance in the fusion process. In other words, we just consider the qualitative counterpart of Example 1 only.

4.1 Fusion of sources using $\mathcal{L}_2$

Example 10: When using $\mathcal{L}_2$, the qualitative inputs[10] of the VBIED problem are chosen according to Table 62.

focal element	$qm_0(.)$	$qm_1(.)$	$qm_2(.)$	$qm_3(.)$
$\theta_4 \cup \theta_8$	L_3	L_0	L_1	L_0
$\theta_6 \cup \theta_8$	L_0	L_2	L_0	L_1
$\overline{\theta_4 \cup \theta_8}$	L_0	L_0	L_2	L_0
I_t	L_0	L_1	L_0	L_2

Table 62: Qualitative inputs using $\mathcal{L}_2$.

Using DSm field and linear algebra of refined labels based on equidistant labels assumption, one gets the following mapping between labels and numbers $L_0 \equiv 0$, $L_1 \equiv 1/3$, $L_2 \equiv 2/3$ and $L_3 \equiv 1$ and therefore, the Table 62 is equivalent to the quantitative inputs table 63 (which are close to the numerical values taken in Example 1).

Applying PCR5 and PCR6 fusion rules, one gets the results given in Table 64 for quantitative and approximate qualitative bba's.

[10] When dealing with qualitative information, we prefix the notations with 'q' letter, for example quantitative bba $m(.)$ becomes qualitative bba $qm(.)$, etc.

focal element	$m_0(.)$	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	1	0	1/3	0
$\theta_6 \cup \theta_8$	0	2/3	0	1/3
$\overline{\theta_4 \cup \theta_8}$	0	0	2/3	0
I_t	0	1/3	0	2/3

Table 63: Corresponding quantitative inputs.

focal element	$m_{PCR5} \approx qm_{PCR5}$	$qm_{PCR6} \approx qm_{PCR6}$
θ_8	$0.25926 \approx L_1$	$0.25926 \approx L_1$
$\theta_4 \cup \theta_8$	$0.36145 \approx L_1$	$0.3157 \approx L_1$
$\theta_6 \cup \theta_8$	$0.093855 \approx L_0$	$0.13198 \approx L_0$
$\overline{\theta_4 \cup \theta_8}$	$0.19158 \approx L_1$	$0.16108 \approx L_0$
I_t	$0.093855 \approx L_0$	$0.13198 \approx L_0$

Table 64: Results of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 62.

One sees that the crude approximation of numerical values to their closest corresponding labels in $\mathcal{L}_2$ can yield to unnormalized qualitative bba. For example, $qm_{PCR6}(.)$ is not normalized since the sum of labels of focal elements in the right column of Table 64 is $L_1 + L_1 + L_0 + L_0 + L_0 = L_2 \neq L_3$. To preserve the normalization of qbba result it is better to work with refined labels as suggested in [6], Vol.3, Chap. 2. Using refined labels, one will get now a better approximation as shown in the Table 65.

focal element	$m_{PCR5} \approx qm_{PCR5}$	$m_{PCR6} \approx qm_{PCR6}$
θ_8	$0.25926 \approx L_{0.79}$	$0.25926 \approx L_{0.79}$
$\theta_4 \cup \theta_8$	$0.36145 \approx L_{1.08}$	$0.31570 \approx L_{0.95}$
$\theta_6 \cup \theta_8$	$0.09385 \approx L_{0.28}$	$0.13198 \approx L_{0.39}$
$\overline{\theta_4 \cup \theta_8}$	$0.19158 \approx L_{0.57}$	$0.16108 \approx L_{0.48}$
I_t	$0.09385 \approx L_{0.28}$	$0.13198 \approx L_{0.39}$

Table 65: Results of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ with refined labels.

It can be easily verified that the qbba's based on refined label approximations are now (qualitatively) normalized (because the sum of refined labels of each column is equal to L_3).

The results of qDSmP based on refined and crude approximations are given in Table 66.

Singletons	$qDSmP_{\epsilon,PCR5}(.)$	$qDSmP_{\epsilon,PCR6}(.)$
θ_1	$0.0323 \approx L_{0.10} \approx L_0$	$0.0273 \approx L_{0.08} \approx L_0$
θ_2	$0.0323 \approx L_{0.10} \approx L_0$	$0.0273 \approx L_{0.08} \approx L_0$
θ_3	$0.0323 \approx L_{0.10} \approx L_0$	$0.0273 \approx L_{0.08} \approx L_0$
θ_4	$0.0017 \approx L_{0.00} \approx L_0$	$0.0017 \approx L_{0.01} \approx L_0$
θ_5	$0.0323 \approx L_{0.10} \approx L_0$	$0.0274 \approx L_{0.08} \approx L_0$
θ_6	$0.0326 \approx L_{0.10} \approx L_0$	$0.0279 \approx L_{0.08} \approx L_0$
θ_7	$0.0323 \approx L_{0.10} \approx L_0$	$0.0273 \approx L_{0.08} \approx L_0$
θ_8	$0.8042 \approx L_{2.40} \approx L_2$	$0.8338 \approx L_{2.51} \approx L_3$

Table 66: Results of $qDSmP_\epsilon$ for Table 62.

Answer to Q1 using crude approximation: Based on these qualitative results, one sees that using crude

approximation (i.e. using only labels in $\mathcal{L}_2$) one gets[11]

- with qPCR5

$$\begin{aligned} qP(\theta_6 \cup \theta_7 \cup \theta_8) &\in [L_1, L_3] \\ qP(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) &\in [L_0, L_2] \\ qP(\theta_7 \cup \theta_8) &\in [L_1, L_3] \\ qP(\overline{\theta_7 \cup \theta_8}) &\in [L_0, L_2] \\ qP(\theta_8) &\in [L_1, L_2] \\ qP(\bar{\theta}_8) &\in [L_1, L_2] \end{aligned}$$

- with qPCR6

$$\begin{aligned} qP(\theta_6 \cup \theta_7 \cup \theta_8) &\in [L_1, L_2] \\ qP(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) &\in [L_1, L_2] \\ qP(\theta_7 \cup \theta_8) &\in [L_1, L_2] \\ qP(\overline{\theta_7 \cup \theta_8}) &\in [L_1, L_2] \\ qP(\theta_8) &\in [L_1, L_2] \\ qP(\bar{\theta}_8) &\in [L_1, L_2] \end{aligned}$$

These results show that is is almost impossible to answer clearly and fairly to the question Q1 using the max of Bel or the max of Pl criteria based on such very inaccurate qualitative bba's using such crude approximation. However it is possible and easy to answer to Q1 using qualitative DSmP value. However and according to Table 66, the final decision must be to **evacuate the building** B when considering the level of DSmP values of $\theta_6 \cup \theta_7 \cup \theta_8$, $\theta_7 \cup \theta_8$, or θ_8.

Answer to Q1 using refined approximation: Using the refined approximation using refined labels which is more accurate, one gets

- with qPCR5

$$\begin{aligned} qP(\theta_6 \cup \theta_7 \cup \theta_8) &\in [L_{1.07}, L_3] \\ qP(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) &\in [L_0, L_{1.93}] \\ qP(\theta_7 \cup \theta_8) &\in [L_{0.79}, L_3] \\ qP(\overline{\theta_7 \cup \theta_8}) &\in [L_0, L_{2.21}] \\ qP(\theta_8) &\in [L_{0.79}, L_{2.43}] \\ qP(\bar{\theta}_8) &\in [L_{0.57}, L_{2.21}] \end{aligned}$$

- with qPCR6

$$\begin{aligned} qP(\theta_6 \cup \theta_7 \cup \theta_8) &\in [L_{1.18}, L_3] \\ qP(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) &\in [L_0, L_{1.82}] \\ qP(\theta_7 \cup \theta_8) &\in [L_{0.79}, L_3] \\ qP(\overline{\theta_7 \cup \theta_8}) &\in [L_0, L_{2.21}] \\ qP(\theta_8) &\in [L_{0.79}, L_{2.52}] \\ qP(\bar{\theta}_8) &\in [L_{0.48}, L_{2.21}] \end{aligned}$$

One sees that accuracy of the result obtained using refined labels allows us to take the decision more easily. Indeed, using the refined approximation, it is possible here to take the decision based on the max of Bel, or on the max of Pl and whatever the decision-support hypothesis used ($\theta_6 \cup \theta_7 \cup \theta_8$, or $\theta_7 \cup \theta_8$, or θ_8), the answer to question Q1 is: **Evacuation of the building** B. The same decision can also be taken from the analysis of qDSmP values as well when considering refined labels in Table 66.

4.2 Fusion of sources using $\mathcal{L}_3$

Here we propose to go further in our analysis and to use a bit more refined set of labels defined by $\mathcal{L}_3$. We need to adapt the qualitative inputs of the VBIED problem in order to work with $\mathcal{L}_3$.

Example 11: We propose to solve the VBIED problem for the following qualitative inputs which reflects what is reported by the sources when using labels belonging to $\mathcal{L}_3$.

focal element	$qm_0(.)$	$qm_1(.)$	$qm_2(.)$	$qm_3(.)$
$\theta_4 \cup \theta_8$	L_4	L_0	L_1	L_0
$\theta_6 \cup \theta_8$	L_0	L_3	L_0	L_1
$\overline{\theta_4 \cup \theta_8}$	L_0	L_0	L_3	L_0
I_t	L_0	L_1	L_0	L_3

Table 67: Qualitative inputs based on $\mathcal{L}_3$.

Based on the equidistant labels assumption, one gets the following mapping between labels and numbers $L_0 \equiv 0$, $L_1 \equiv 1/4$, $L_2 \equiv 2/4$, $L_3 \equiv 3/4$ and $L_4 = 1$ and therefore, the Table 67 is equivalent to the quantitative inputs table 68 (which are more close to the numerical values taken in Example 1 than the inputs chosen in Table 63 for Example 10).

focal element	$m_0(.)$	$m_1(.)$	$m_2(.)$	$m_3(.)$
$\theta_4 \cup \theta_8$	1	0	0.25	0
$\theta_6 \cup \theta_8$	0	0.75	0	0.25
$\overline{\theta_4 \cup \theta_8}$	0	0	0.75	0
I_t	0	0.25	0	0.75

Table 68: Corresponding quantitative inputs.

Applying PCR5 and PCR6 fusion rules, one gets the results given in Table 69 for quantitative and approxi-

[11] The derivations of $qBel(\bar{X})$ and $qPl(\bar{X})$ were obtained using qualitative extension of Dempster's formulas [5], i.e. with $qBel(\bar{X}) = L_m - qPl(X)$ and $qPl(\bar{X}) = L_m - qBel(X)$. These results are valid only if the qbba is normalized, but are used here even when using non normalized qbba as crude approximation.

mate qualitative bba's using refined and crude approximations of labels.

foc. elem.	$m_{PCR5} \approx qm_{PCR5}$	$m_{PCR6} \approx qm_{PCR6}$
θ_8	$0.20312 \approx L_{0.81} \approx L_1$	$0.20312 \approx L_{0.81} \approx L_1$
$\theta_4 \cup \theta_8$	$0.34269 \approx L_{1.37} \approx L_1$	$0.29979 \approx L_{1.21} \approx L_1$
$\theta_6 \cup \theta_8$	$0.11617 \approx L_{0.47} \approx L_0$	$0.15370 \approx L_{0.61} \approx L_1$
$\overline{\theta_4 \cup \theta_8}$	$0.22185 \approx L_{0.88} \approx L_1$	$0.18969 \approx L_{0.76} \approx L_1$
I_t	$0.11617 \approx L_{0.47} \approx L_0$	$0.15370 \approx L_{0.61} \approx L_1$

Table 69: Results of $m_0 \oplus m_1 \oplus m_2 \oplus m_3$ for Table 67.

It can be easily verified that the qbba's based on refined label approximations are (qualitatively) normalized because the sum of refined labels of each column is equal to L_4. Using crude approximation when working only with labels in $\mathcal{L}_3$ we get non normalized qbba's. The results of qDSmP based on refined and crude approximations are given in Table 70.

Singletons	$qDSmP_{\epsilon,PCR5}(.)$	$qDSmP_{\epsilon,PCR6}(.)$
θ_1	$0.0375 \approx L_{0.15} \approx L_0$	$0.0323 \approx L_{0.13} \approx L_0$
θ_2	$0.0375 \approx L_{0.15} \approx L_0$	$0.0323 \approx L_{0.13} \approx L_0$
θ_3	$0.0375 \approx L_{0.15} \approx L_0$	$0.0323 \approx L_{0.13} \approx L_0$
θ_4	$0.0022 \approx L_{0.01} \approx L_0$	$0.0022 \approx L_{0.01} \approx L_0$
θ_5	$0.0375 \approx L_{0.15} \approx L_0$	$0.0323 \approx L_{0.13} \approx L_0$
θ_6	$0.0381 \approx L_{0.15} \approx L_0$	$0.0331 \approx L_{0.13} \approx L_0$
θ_7	$0.0375 \approx L_{0.15} \approx L_0$	$0.0323 \approx L_{0.13} \approx L_0$
θ_8	$0.7722 \approx L_{3.09} \approx L_3$	$0.8032 \approx L_{3.21} \approx L_3$

Table 70: Results obtained with $qDSmP_\epsilon$ for Table 67.

Ones sees that the use of refined labels allows to obtain normalized qualitative probabilities. This is not possible to get normalized qualitative probabilities when using only crude approximations with labels in $\mathcal{L}_3$ for this example.

Answer to Q1: Using refined labels (which is more accurate), one gets finally

- with qPCR5

$$\begin{aligned} qP(\theta_6 \cup \theta_7 \cup \theta_8) &\in [L_{1.28}, L_4] \\ qP(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) &\in [L_0, L_{2.72}] \\ qP(\theta_7 \cup \theta_8) &\in [L_{0.81}, L_4] \\ qP(\overline{\theta_7 \cup \theta_8}) &\in [L_0, L_{3.19}] \\ qP(\theta_8) &\in [L_{0.81}, L_{3.12}] \\ qP(\bar{\theta}_8) &\in [L_{0.88}, L_{3.19}] \end{aligned}$$

- with qPCR6

$$\begin{aligned} qP(\theta_6 \cup \theta_7 \cup \theta_8) &\in [L_{1.42}, L_4] \\ qP(\overline{\theta_6 \cup \theta_7 \cup \theta_8}) &\in [L_0, L_{2.58}] \\ qP(\theta_7 \cup \theta_8) &\in [L_{0.81}, L_4] \\ qP(\overline{\theta_7 \cup \theta_8}) &\in [L_0, L_{3.19}] \\ qP(\theta_8) &\in [L_{0.81}, L_{3.24}] \\ qP(\bar{\theta}_8) &\in [L_{0.76}, L_{3.19}] \end{aligned}$$

One sees that based with PCR5 or PCR6 whatever the decision-support hypothesis we consider ($\theta_6 \cup \theta_7 \cup \theta_8$, $\theta_7 \cup \theta_8$, or θ_8), one will decide to evacuate the building B based on max of Bel, max of Pl or DSmP values, except for the case of PCR5 with θ_8 based on the max of Bel, max of Pl. In this case, PCR5 result suggests to NOT evacuate B contrariwise to PCR6 result. As far as θ_8 is the preferred (optimistic) decision-support hypothesis, one sees here the main effect of difference between PCR5 and PCR6 for decision-making support. But as already stated, for such problem the most prudent strategy for decision-making is to consider the decisio-support hypothesis $\theta_6 \cup \theta_7 \cup \theta_8$ which captures all aspects of potential danger. Using such reasonable strategy, both rules PCR5 and PCR6 yields same decision: Evacuation of the building B.

5 Conclusions

In this paper we have presented a modeling for solving the Vehicle-Born Improvised Explosive Device (VBIED) problem with Dezert-Smarandache Theory (DSmT) framework. We have shown how it is possible to compute imprecise probabilities of all decision-support hypotheses and how to take into account the reliabilities and the importances of the sources of information in decision-making support. The strong impact of prior information has also been analyzed, as well as the possibility to deal directly with imprecise sources of information and even with qualitative reports. We have answered with the full justification to the two main questions asked in the VBIED problem by John Lavery and Simon Maskell: 1) what is the final decision to take, and 2) what is the best fusion subsystem to choose (APNR or the pool of experts)? The analysis done in this paper is based on a very limited number of reasonable assumptions and could be adapted for solving more complicated security problems involving imprecise, incomplete and conflicting sources of information.

References

[1] F. Dambreville, *Generic implementation of fusion rules based on Referee function*, Proc. of Workshop on the theory of belief functions, April 1-2, 2010, Brest, France.

[2] J. Dezert, J.-M. Tacnet, M. Batton-Hubert, F. Smarandache, *Multi-criteria decision making based on DSmT-AHP*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[3] A. Martin, C. Osswald, *A new generalization of the proportional conflict redistribution rule stable in terms of decision*, Chapter 2 in [6], Vol. 2, 2006.

[4] A. Martin, *Implementing general belief function framework with a practical codification for low complexity*, Chapter 7 in [6], Vol. 3, 2009.

[5] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[6] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3, American Research Press, 2004–2009. http://fs.gallup.unm.edu/DSmT.htm

[7] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, accepted in Fusion 2010 conf, Edinburgh, July 26-29, 2010.

[8] Ph. Smets Ph., *The Combination of Evidence in the Transferable Belief Model*, IEEE Trans. PAMI 12, pp. 447–458, 1990.

[9] J. Sudano, *The system probability information content (PIC) relationship to contributing components, combining independent multi-source beliefs, hybrid and pedigree pignistic probabilities*, Proc. of Fusion 2002, Vol.2, pp. 1277-1283, Annapolis, MD, USA, July 2002.

A PCR-BIMM filter For Maneuvering Target Tracking

Jean Dezert
Benjamin Pannetier

Originally published as Dezert J., Pannetier B., *A PCR-BIMM filter for maneuvering target tracking,* in Proc. of Fusion 2010, Edinburgh, Scotland, UK, 26-29 July 2010, and reprinted with permission.

Abstract – *In this paper we show how to correct and improve the Belief Interacting Multiple Model filter (BIMM) proposed in 2009 by Nassreddine et al. for tracking maneuvering targets. Our improved algorithm, called PCR-BIMM is based on results developed in DSmT (Dezert-Smarandache Theory) framework and concerns two main steps of BIMM: 1) the update of the basic belief assignment of modes which is done by the Proportional Conflict Redistribution Rule no. 5 rather than Smets' rule (conjunctive rule); 2) the global target state estimation which is obtained from the DSmP probabilistic transformation rather than the commonly used Pignistic transformation. Monte-Carlo simulation results are presented to show the performances of this PCR-BIMM filter with respect to classical IMM and BIMM filters obtained on a very simple maneuvering target tracking scenario.*

Keywords: Tracking, Maneuvering target, IMM, BIMM, DSmT.

1 Introduction

In Fusion 2009 international conference, Nassreddine, Abdallah, and Denœux [13] have proposed an interesting idea to extend the classical Interacting Multiple Models (IMM) filter with belief function theory in order to deal with an unknown and variant motion models. Their algorithm is based on the classical/historical belief function theory developed by Shafer in 1976 [14], known as Dempster-Shafer Theory (DST) and requires both Smets' rule, i.e. the conjunctive fusion rule equivalent to the non normalized Dempster's rule, and the probabilistic pignistic transformation. This algorithm is called Belief Interacting Multiple Model algorithm (BIMM). According to authors results, BIMM algorithm outperforms classical IMM algorithm at least in the vehicle localization problem studied in their works. These appealing results and the possible extension of IMM in belief function theory framework motivates our interest to analyze and evaluate this new BIMM filter. A deep analysis of the paper yields to the following comments:

1. The derivation of the predicted prior basic belief assignment of modes in Step 1 of BIMM algorithm was clearly wrong in [13] as proved in the sequel. This mistake implies a serious doubt on the validity of the results presented in [13].

2. The simulations results presented in [13] cannot be verified precisely, nor reproduced, because some settings parameters (like α_i discounting factors) required for the BIMM filter have not be provided by the authors and the essential step 9 of the algorithm was not detailed enough.

3. It is known (see Chapter 1 of [15] Vol. 3) that the conjunctive rule does not perform efficiently in a sequential fusion process because the empty set is an absorbing element for the conjunctive fusion rule. Therefore, in order to implement successfully the BIMM filter, some ad-hoc numerical techniques are necessary (or some extra normalization steps) in the BIMM algorithm in order to prevent the mass of belief committed to empty set to become close to one and make Smets' rule responding to new information. This serious problem has unfortunately not been discussed in [13].

From the theoretical point of view, it is quite surprising that one gets better performances with the BIMM (which proceeds with less specific information since it deals with non Bayesian basic belief assignments) than with the classical Bayesian IMM filter (which deals with more specific information, i.e. with Bayesian basic belief assignments). The first purpose of this work is to verify if the conclusions given in [13] are valid on a very simple reproducing maneuvering target tracking scenario. We want also to see if a more justified Belief-based IMM algorithm can be developed to improve the BIMM algorithm and to evaluate it to get a fair comparison of its performance with respect to

classical IMM filter. The improvement of the BIMM algorithm we propose in this paper is based on advanced theoretical results obtained in the development of Dezert-Smarandache Theory (DSmT) of information fusion [15]. This paper is organized as follows: After a brief recall of classical (fixed structure) IMM algorithm given in section 2, one presents in section 3 the Belief IMM algorithm and its flaws. Motivations for the improvement of the BIMM filter is presented in section 4 with the presentation of the main steps of our new algorithm called PCR-BIMM filter (Proportional Conflict Redistribution-based BIMM). In section 5, we examine the performances of the IMM, and PCR-BIMM on a very simple tracking scenario through Monte-Carlo results. Conclusions and perspectives for further investigations are given in section 6.

2 Classical IMM algorithm

The IMM filter is one of the most used algorithm for tracking maneuvering targets and was developed originally by Henk Blom in eighties [5, 6, 2]. The IMM filter is a recursive filter with a low complexity and has been proved very efficient in many real-data tracking applications [4] and many extensions of IMM have been developed since its original publication for dealing with multitarget-multisensor case, cluttered environments, etc, see [12] for a good survey of Multiple Models techniques. The classical IMM algorithm considers a hybrid Multiple Models (MM) system which obeys one of a finite number r of dynamic models M_i, $i = 1, \ldots, r$ and estimates the posterior mode probabilities from their prior probabilities and target measurements (Bayesian framework). Its specificity is that IMM mixes hypotheses with depth 1 only at the start of each cycle and thus has a low complexity of order $O(r)$, while providing same performances as the more effective Generalized Pseudo-Bayesian estimator of order 2. We briefly recall the principle of classical IMM filter, see [3, 4] for more details with examples. A hybrid MM system is characterized by two state variables: 1) the *base-state* variable $\mathbf{x}(k)$ of dimension n_x including the position, velocity, etc. of the target, and 2) a *modal-state* $M_j(k)$ belonging to a known finite set $\mathcal{M}_r(k) = \{M_i(k), i = 1, \ldots, r\}$ of r possible dynamic models for the target during its motion. For simplicity of presentation, we consider only a fixed-structure IMM, i.e. $\mathcal{M}_r(k) = \mathcal{M}_r$ is invariant with time. Variable-structure IMM is possible and has been introduced by Xiao-Rong Li in [10, 11]. The hybrid system is described by the equations[1]

$$\mathbf{x}(k) = \mathbf{F}[M(k)]\mathbf{x}(k-1) + \mathbf{v}[k-1, M(k)]$$

$$\mathbf{z}(k) = \mathbf{H}[M(k)]\mathbf{x}(k) + \mathbf{w}[k, M(k)]$$

where $M(k)$ is the mode in effect during the sampling period ending at time k belonging in $\mathcal{M}_r$. $\mathbf{x}(k)$ and $\mathbf{z}(k)$ are the target state and observation vectors. The set of all available measurements up to k is denoted $\mathbf{Z}^k$. $\mathbf{F}[M(k)]$ and $\mathbf{H}[M(k)]$ are known matrices depending on the dynamic model $M(k)$. The statistics of the process and observation noises $\mathbf{v}[k-1, M(k)]$ and $\mathbf{w}[k, M(k)]$ can differ from mode to mode. Usually one considers $\mathbf{v}[k-1, M(k) = M_j] \sim \mathcal{N}(\bar{\mathbf{v}}_j, \mathbf{Q}_j)$ and $\mathbf{w}[k, M(k) = M_j] \sim \mathcal{N}(\bar{\mathbf{w}}_j, \mathbf{R}_j)$ with known covariance matrices $\mathbf{Q}_j$ and $\mathbf{R}_j$ respectively. The Mode jump process is modeled as a Makov chain with known a priori probabilities $P\{M(0) = M_j\} = \mu_j(k=0)$ and known transition probabilities $P\{M(k) = M_j | M(k-1) = M_i\} = \pi_{ij}$. A cycle of the classical IMM algorithm $(k-1) \mapsto k$ consists in the following steps:

• **Step 0** (Initialization at $k = 0$): Definition of dynamic and observation matrices, choice of process and observation noise levels, sampling period, initialization of the filters adapted to each mode, choice of the prior mode probabilities P_j and the transition probability matrix $\mathbf{P}_t \triangleq [\pi_{ij} = P\{M_j(k)|M_i(k-1)\}]'$ assumed known and time-invariant.

• **Step 1** (Interaction-mixing $(j = 1, \ldots, r)$): Mixing of the previous cycle mode-conditioned state estimates $\hat{\mathbf{x}}_i(k-1|k-1)$ and covariance, using the mixing probabilities $\mu_{i|j}(k-1|k-1)$, to initialize the current cycle of each mode-conditioned filter $\hat{\mathbf{x}}^0_j(k-1|k-1)$. This is done by

$$\hat{\mathbf{x}}^0_j(k-1|k-1) = \sum_{i=1}^{r} \mu_{i|j}(k-1|k-1)\hat{\mathbf{x}}_i(k-1|k-1) \quad (1)$$

$$\begin{aligned}\mathbf{P}^0_j(k-1|k-1) = \sum_{i=1}^{r} \mu_{i|j}(k-1|k-1)\{&\mathbf{P}_i(k-1|k-1) \\ &- [\hat{\mathbf{x}}_i(k-1|k-1) - \hat{\mathbf{x}}^0_j(k-1|k-1)]\cdot \\ &[\hat{\mathbf{x}}_i(k-1|k-1) - \hat{\mathbf{x}}^0_j(k-1|k-1)]'\} \quad (2)\end{aligned}$$

where the elements $\mu_{i|j}(k-1|k-1)$ of the mixing probability (vertical) vector $\boldsymbol{\mu}_{k-1|k-1}(.|M_j(k)) = [\mu_{i|j}(k-1|k-1), i = 1, \ldots r]'$ are calculated by

$$\begin{aligned}\mu_{i|j}(k-1|k-1) &\triangleq P\{M_i(k-1)|M_j(k), \mathbf{Z}^{k-1}\} \\ &= \frac{\pi_{ij}\mu_i(k-1)}{\mu^-_j(k)} \quad (3)\end{aligned}$$

with

$$\mu^-_j(k) \triangleq P\{M_j(k)|\mathbf{Z}^{k-1}\} = \sum_{i=1}^{r} \pi_{ij}\mu_i(k-1) \quad (4)$$

The equation (4) can be written more concisely as:

$$\boldsymbol{\mu}^-_k(.) = \mathbf{P}_t \cdot \boldsymbol{\mu}_{k-1}(.) \quad (5)$$

where $\mathbf{P}'_t = [\pi_{ij}]$ and $\boldsymbol{\mu}_{k-1}(.)$ represents the (vertical) vector of prior probability of modes, i.e.

$$\boldsymbol{\mu}_{k-1}(.) = [P(M_i(k-1)|\mathbf{Z}^{k-1}]' = [\mu_1(k-1) \ \ldots \mu_r(k-1)]'$$

[1] For simplicity, we assume here linear systems.

and $\boldsymbol{\mu}_k^-(.)$ represents the (vertical) vector of predicted prior probability of modes

$$\boldsymbol{\mu}_k^-(.) = [P(M_j(k)|\mathbf{Z}^{k-1}]' = [\mu_1^-(k) \ \ldots \mu_r^-(k)]'$$

• **Step 2** (Mode conditioned filter): From prior mixed statistics $\hat{\mathbf{x}}_j^0(k-1|k-1)$ and $\mathbf{P}_j^0(k-1|k-1)$ and the target measurement $\mathbf{z}(k)$, one calculates $\hat{\mathbf{x}}_j(k|k)$ and $\hat{\mathbf{P}}_j(k|k)$ for each possible mode in effect (r filters running in parallel) by a specific filter matched to mode M_j, typically a Kalman filter if the dynamic and observation system are linear, or Extended Kalman Filter (EKF) to deal with linear or non linear equations, or any other sophisticated filters if necessary for dealing for example with miss-detections and false alarms [3]. The likelihood $\Lambda_j(k)$ of the filter j is assumed to be Gaussian with

$$\Lambda_j(k) = \frac{1}{(2\pi)^{n_z/2}\sqrt{|\mathbf{S}_j(k)|}} \exp^{-\frac{1}{2}\tilde{\mathbf{z}}_j'(k)\mathbf{S}_j^{-1}(k)\tilde{\mathbf{z}}_j(k)} \quad (6)$$

where $\tilde{\mathbf{z}}_j(k) \triangleq \mathbf{z}(k) - \hat{\mathbf{z}}_j(k|k-1)$ is the innovation and $\mathbf{S}_j(k)$ is the covariance of the innovation provided by the filter j.

• **Step 3** (Mode probability update): The probability $\mu_j(k)$ of each mode j for $j = 1, \ldots, r$ is calculated by

$$\mu_j(k) = P\{M_j(k)|\mathbf{Z}^k\} = \Lambda_j(k)\mu_j^-(k) / \sum_{i=1}^{r} \Lambda_i(k)\mu_i^-(k) \quad (7)$$

• **Step 4** (Global estimation for output purpose): The global estimate $\hat{\mathbf{x}}(k|k)$ and the covariance of estimation error $\mathbf{P}(k|k)$ are given by:

$$\hat{\mathbf{x}}(k|k) = \sum_{j=1}^{r} \mu_j(k)\hat{\mathbf{x}}_j(k|k) \quad (8)$$

$$\mathbf{P}(k|k) = \sum_{j=1}^{r} \mu_j(k)\{\mathbf{P}_j(k|k) - [\hat{\mathbf{x}}_j(k|k) - \hat{\mathbf{x}}(k|k)] \cdot [\hat{\mathbf{x}}_j(k|k) - \hat{\mathbf{x}}(k|k)]'\} \quad (9)$$

3 Belief-based IMM algorithm

In 2009, Nassreddine et al. have proposed in [13] an extension of classical IMM filter in the framework of Dempster-Shafer Theory (DST) [14] for dealing with an unknown and variant motion models. The idea was to select a set of candidate models[2], and then estimate a current basic belief assignment (bba) defined on the power-set of this set of models based on the fusion of bba's built from measurement likelihoods with the predicted bba of the models using Smets' rule[3] denoted $\textcircled{\scriptsize}$. From the result of Smets' fusion, the mixed state of classical IMM filter is replaced with the *pignistic averaging* of the mode-conditioned state estimates. This new extension of IMM filter was called BIMM (Belief-based IMM) since it uses belief function theory to represent the uncertainty in the switches between the modes. This section presents succinctly the principle of the BIMM filter. We justify also our motivation for developing a new Belief-based IMM algorithm. The steps of BIMM are actually very close to the steps of classical IMM, except that predicted and updated mode probabilities are estimated from pignistic probabilities derived from a basic belief assignment updated with the conjunctive rule of combination. The main changes of BIMM concern the Step 1 and the Step 3 of IMM algorithm. The frame of discernment chosen in BIMM coincides with the set of possible models, i.e. $\Theta(k) \equiv \mathcal{M}_r(k) = \{M_i(k), i = 1, \ldots, r\}$. Instead of computing recursively the mixed $\mu_{i|j}(.)$ and updated $\mu_j(.)$ probabilities with eqs. (3) and (4) as done with the classical IMM, one deals with bba's defined on the power-set 2^Θ of the frame of discernment. Mathematically, a normal bba $\mathbf{m}(.)$ is defined[4] as a mapping from $2^\Theta \mapsto [0,1]$ such that $m(\emptyset) = 0$ and $\sum_{A\in 2^\Theta} m(A) = 1$. A is a focal element of $\mathbf{m}(.)$ if $m(A) > 0$. Any discrete probability measure can be interpreted as a special belief function, called Bayesian belief [14] whose focal elements are singletons of 2^Θ. Any belief function with a bba $\mathbf{m}(.)$ can be approximated into subjective probability measure thanks to the pignistic transformation [17] defined for all $M_i \in \Theta(k)$ by

$$BetP\{M_i\} = \sum_{A\in 2^\Theta | A\cap M_i = M_i} \frac{1}{|A|} \cdot \frac{m(A)}{1 - m(\emptyset)} \quad (10)$$

where $|A|$ is the cardinality of A.

The steps of BIMM proposed in [13] are[5]:

• **Step 0** (Initialization at $k = 0$): Definition of dynamic and observation matrices, choice of process and observation noise levels, sampling period, initialization of the filters adapted to each mode. The prior probabilities of modes $\{P_j = P\{M(0) = M_j\}, j = 1, \ldots, r\}$ used in IMM, are replaced[6] by the vacuous belief assignment $m(\Theta(k=0) = M_1 \cup M_2 \cup \ldots \cup M_r) = 1$. The probability transition matrix $\mathbf{P}_t' = [\pi_{ij}]$ is replaced by a bba transition matrix[7] $\mathbf{M}_t \triangleq [m_{ij}]$ having a very simple structure defined by the r implication rules: "R_i: if $M(k) = M_i(k)$ then $M(k+1) = M_i(k+1)$" with known belief coefficients $\beta_i \in [0,1]$ for $i = 1, 2, \ldots, r$ with $\beta_i = m(M_i(k+1)|M_i(k))$ and $1 - \beta_i = m(\Theta(k+1) = M_1 \cup \ldots \cup M_r(k+1)|M_i(k))$.

[2] Corresponding to the so-called frame of discernment and usually denoted Θ in DST.

[3] Smet's rule is nothing but the non normalized Dempster's rule of combination, i.e. the conjunctive rule.

[4] We use boldface letters to denote vectors or matrices.

[5] We use a more classical notation generally adopted in the tracking community.

[6] Note that this initialization can also be done by taking $m(M_j(k=0) = P_j$ as well if one considers that prior probabilities of modes is accurate enough.

[7] Called switching mass function in [13].

• **Step 1** (Interaction-mixing): The mixing probability $\mu_{i|j}(k-1|k-1)$ are calculated as follows:

1. The derivation of probabilities vector $\boldsymbol{\mu}_k^-(.) = [\mu_1^-(k) \ \ldots \ \mu_r^-(k)]'$ in classical IMM is replaced by the derivation of the predicted bba $\mathbf{m}_k^-(.)$ given by

$$\mathbf{m}_k^-(.) \triangleq \mathbf{M}_t \cdot \mathbf{m}_{k-1}(.) \tag{11}$$

2. The derivation of probabilities $\mu_{i|j}(k-1|k-1) \triangleq P\{M_i(k-1)|M_j(k), \mathbf{Z}^{k-1}\}$ is replaced by the derivation of bba $m_{k-1|k-1}(.)$ thanks to the Generalized Bayesian Theorem (GBT) [18]. More precisely,

$$\begin{aligned}&\mathbf{m}_{k-1|k-1}(.|M_j(k)) = \\ &[\circledcirc\mathbf{m}_k^{\Uparrow\Theta(k-1)\times\Theta(k)}(.|M_i(k-1))](.|M_j(k)]^{\downarrow\Theta(k-1)}\end{aligned} \tag{12}$$

where $\Uparrow \Theta(k-1)\times\Theta(k)$ is the ballooning extension [18] of the bba on the Cartesian product frame $\Theta(k-1)\times\Theta(k)$, and where $\downarrow\Theta(k-1)$ represents the marginalization operation of the bba on the frame $\Theta(k-1)$. See [18], for details and examples.

3. The derivation of the mixing probability $\mu_{i|j}(k-1|k-1) = P\{M_i(k-1)|M_j(k), \mathbf{Z}^{k-1}\}$ of classical IMM is replaced by the pignistic probability drawn from $\mathbf{m}_{k-1|k-1}(.|M_j(k))$, that is:

$$\mu_{i|j}(k-1|k-1) = BetP\{M_i(k-1)|M_j(k), \mathbf{Z}^{k-1}\}$$

where $BetP\{.\}$ is calculated with the transformation (10) using $\mathbf{m}_{k-1|k-1}(.|M_j(k))$ given by (12).

$\hat{\mathbf{x}}_j^0(k-1|k-1)$ and $\mathbf{P}_j^0(k-1|k-1)$ are calculated as in IMM Step 1.

• **Step 2**: Same as IMM Step 2.

• **Step 3** (Mode bba update): The updated bba $\mathbf{m}_k(.)$ of modes is computed from the conjunctive combination of the predicted bba $\mathbf{m}_{k-1}^-(.)$ given in (11) with *observed bba's*[8] $\mathbf{m}_{k,j}(.)$, $j = 1, 2, \ldots r$ by

$$\mathbf{m}_k(.) = [\mathbf{m}_{k,1}\circledcirc \ldots \circledcirc\mathbf{m}_{k,r}\circledcirc\mathbf{m}_{k-1}^-](.) \tag{13}$$

where the *observed bba's* $\mathbf{m}_{k,j}(.)$ for $j = 1, \ldots, r$ are given[9] by [13]:

$$\begin{cases} m_{k,j}(M_j(k)) & = 0 \\ m_{k,j}(\bar{M}_j(k)) & = \alpha_j(1 - R\Lambda_j(k)) \\ m_{k,j}(\Theta(k)) & = 1 - \alpha_j(1 - R\Lambda_j(k)) \end{cases} \tag{14}$$

α_j is a discounting coefficient associated with the likelihood of the mode $M_j(k)$ and R is a normalization constant.

[8] We mean that the bba $\mathbf{m}_{k,j}(.)$ is built from the likelihood $\Lambda_j(k)$ which depends on the mode $M_j(k)$ and on the observation available $\mathbf{z}(k)$.

[9] This is Appriou's model no. 1 in [1].

• **Step 4** (Global estimation for output purpose): The global estimate $\hat{\mathbf{x}}(k|k)$ and the covariance of estimation error $\mathbf{P}(k|k)$ are given as in step 4 of classical IMM by taking $\mu_j(k) = BetP\{M_j(k)|\mathbf{Z}^k\}$ where $BetP\{M_j(k)|\mathbf{Z}^k\}$ is the pignistic probability that the mode M_j is effective at time k. $BetP\{M_j(k)|\mathbf{Z}^k\}$ is computed from the updated bba $\mathbf{m}_k(.)$ given by (13).

A mistake in Step 1 of BIMM filter: The aforementioned Step 1 of BIMM algorithm described with an example in [13] is clearly incorrect because the derivation of the predicted bba $\mathbf{m}_k^-(.)$ by (5) is wrong because the sum of masses of focal elements is not equal to one. It is easy to verify from example in [13] when considering only two models, when taking $\beta_1 = m_(M_1(k)|M_1(k-1)) = 0.9$, $1-\beta_1 = 0.1 = m_(M_1(k)\cup\bar{M}_1(k)|M_1(k-1))$ and $\beta_2 = m_(M_2(k)|M_2(k-1)) = 0.89$, $1-\beta_2 = 0.11 = m_(M_2(k)\cup\bar{M}_2(k)|M_2(k-1))$ and taking the prior bba $\mathbf{m}_{k-1}(.) = [m(\emptyset) = 0 \ m(M_1(k-1)) = 0.45 \ m(M_2(k-1)) = 0.20 \ m(M_1(k-1)\cup M_2(k-1)) = 0.35]'$. Applying the wrong formula (11), one gets precisely:

$$\underbrace{\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0.9 & 0 & 0.1\\ 0 & 0 & 0.89 & 0.11\\ 0 & 0 & 0 & 1\end{bmatrix}}_{\mathbf{M}_t}\underbrace{\begin{bmatrix}0\\0.45\\0.20\\0.35\end{bmatrix}}_{\mathbf{m}_{k-1}(.)} = \underbrace{\begin{bmatrix}0\\0.4400\\0.2165\\0.3500\end{bmatrix}}_{\mathbf{m}_k^-(.)} \neq \underbrace{\begin{bmatrix}0\\0.44\\0.21\\0.35\end{bmatrix}}_{\text{Result in [13]}}$$

One can see that the sum of components of $\mathbf{m}_k^-(.)$ equals 1.0065 !!! This mistake is not due to rounding approximation of the result, but to a more serious mistake in the choice of the transition matrix $\mathbf{M}_t$. This mistake actually comes from the confusion in indices of the classical IMM transition matrix. It is easy to verify that the correct transition matrix must be actually taken as the transpose of $\mathbf{M}_t$. Therefore, the correct derivation of $\mathbf{m}_k^-(.)$ must be done by

$$\mathbf{m}_k^-(.) \triangleq \mathbf{M}_t' \cdot \mathbf{m}_{k-1}(.) \tag{15}$$

For the example 1 of [13], one will get correctly

$$\underbrace{\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 0.9 & 0 & 0\\ 0 & 0 & 0.89 & 0\\ 0 & 0.1 & 0.11 & 1\end{bmatrix}}_{\mathbf{M}_t'}\underbrace{\begin{bmatrix}0\\0.45\\0.20\\0.35\end{bmatrix}}_{\mathbf{m}_{k-1}(.)} = \underbrace{\begin{bmatrix}0\\0.4050\\0.1780\\0.4170\end{bmatrix}}_{\mathbf{m}_k^-(.)}$$

Remarks on BIMM filter: The BIMM is based on two[10] pillars: 1) the conjunctive rule of combination, and 2) the pignistic transformation to approximate a bba into a subjective probability measure because. These two pillars are disputable because:

[10] Actually, Smets' Generalized Bayesian Theorem (GBT) could be also considered as the third pillar of BIMM.

1. The efficiency of Smets' rule for combining bba's is very questionable in this belief-based extension of IMM because it has been already proved in [15], Vol. 3, and specially in sequential Target Type Tracking problem [7] that such rule doesn't perform well in general for mode change detection. Smets' rule doesn't respond to new information since very quickly all the mass of belief concentrates on the empty set. See example in [15], Vol. 3, Chap. 1, freely downloadable from the web and not reported here due to space limitation.

2. The real interest and efficiency of the pignistic transformation is also disputable because there exists other probabilistic transformations which perform better than $BetP$ in term of probabilistic informational content, in particular the DSmP transformation developed in [15], Vol. 3, Chap 1 & 3 and also in [8].

3. The justification for the use of Appriou's model no. 1 in step 3 of BIMM is missing and probably other (and maybe better) models could be developed to derive the updated bba $\mathbf{m}_k(.)$. This question has not been investigated in this paper and will be a source for future research.

Interest of BIMM w.r.t. IMM: The potential advantage of the belief-based IMM approach is to offer some robustness of the filter when replacing the strong constraint on the knowledge of probability of transitions π_{ij} (usually based on ad-hoc assumptions on the mean sojourn time of the target in each mode) by a more flexible constraint on the transitions based on (very simple and less specific) uncertain implication rules. With BIMM, one can also relax the knowledge of the prior probabilities of the modes by starting the tracking directly with a vacuous belief prior of the modes. Of course, if one has good reasons to use a given prior of modes, this can be done easily in belief-based IMM approach which is also a nice features of such filter.

4 PCR-BIMM algorithm

To preserve the potential advantages of BIMM and to overcome its aforementionned problems, we propose to keep its general structure as a belief-based extension of classical IMM but we replace Smets' rule by the more effective Proportional Conflict Redistribution rule no. 5 (PCR5), or eventually the more simple PCR rule no. 6 (PCR6), and to replace the pignistic transformation by the more effective DSmP transformation to estimate modes probabilities required in the IMM filter. We call this new algorithm, the PCR-BIMM filter. Before giving the sketch of our PCR-BIMM filter, we just recall what are the PCR5 fusion rule and the DSmP transformation. All details, justifications with examples on PCR5 and DSmP can be found freely from the web in [15], Vols. 2 & 3 and will not be reported here.

4.1 PCR5 and PCR6 fusion rules

In DSmT (Dezert-Smarandache Theory) framework, the Proportional Conflict Redistribution Rule no. 5 (PCR5) is used generally to combine bba's. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. Let $m_1(.)$ and $m_2(.)$ be two independent[11] bba's, then the PCR5 rule is defined as follows (see [15], Vol. 2 for full justification and examples): $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)} \right] \quad (16)$$

where all denominators in (16) are different from zero. If a denominator is zero, that fraction is discarded. Additional properties of PCR5 can be found in [9]. Extension of PCR5 for combining qualitative bba's can be found in [15], Vol. 2 & 3. All propositions/sets are in a canonical form. A variant of PCR5, called PCR6 has been proposed by Martin and Osswald in [15], Vol. 2, for combining $s > 2$ sources. The general formulas for PCR5 and PCR6 rules are given in [15], Vol. 2 also. PCR6 coincides with PCR5 when one combines two sources. The difference between PCR5 and PCR6 lies in the way the proportional conflict redistribution is done as soon as three or more sources are involved in the fusion. For example, let's consider three sources with bba's $m_1(.)$, $m_2(.)$ and $m_3(.)$, $A \cap B = \emptyset$ for the model of the frame Θ, and $m_1(A) = 0.6$, $m_2(B) = 0.3$, $m_3(B) = 0.1$. With PCR5 the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR5} = 0.01714$ and $x_B^{PCR5} = 0.00086$ because the proportionalization requires

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B)m_3(B)}$$

$$\text{that is} \quad \frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{0.03} = \frac{0.018}{0.6 + 0.03} \approx 0.02857$$

$$\text{thus} \quad \begin{cases} x_A^{PCR5} = 0.60 \cdot 0.02857 \approx 0.01714 \\ x_B^{PCR5} = 0.03 \cdot 0.02857 \approx 0.00086 \end{cases}$$

With the PCR6 fusion rule, the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR6} = 0.0108$ and

[11] I.e. each source provides its bba independently of the other sources.

$x_B^{PCR6} = 0.0072$ because the PCR6 proportionalization is done as follows:

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_{B,2}^{PCR6}}{m_2(B)} = \frac{x_{B,3}^{PCR6}}{m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A)+m_2(B)+m_3(B)}$$

that is

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_{B,2}^{PCR6}}{0.3} = \frac{x_{B,3}^{PCR6}}{0.1} = \frac{0.018}{0.6+0.3+0.1} = 0.018$$

thus

$$\begin{cases} x_A^{PCR6} = 0.6 \cdot 0.018 = 0.0108 \\ x_{B,2}^{PCR6} = 0.3 \cdot 0.018 = 0.0054 \\ x_{B,3}^{PCR6} = 0.1 \cdot 0.018 = 0.0018 \end{cases}$$

and therefore with PCR6, one gets finally the following redistributions to A and B:

$$\begin{cases} x_A^{PCR6} = 0.0108 \\ x_B^{PCR6} = x_{B,2}^{PCR6} + x_{B,3}^{PCR6} = 0.0054 + 0.0018 = 0.0072 \end{cases}$$

From the implementation point of view, PCR6 is much more simple to implement than PCR5. For convenience, Matlab codes of PCR5 and PCR6 fusion rules can be found in [15, 16].

4.2 The DSmP transformation

The DSmP probabilistic transformation is a serious alternative to the classical pignistic transformation which allows to increase the probabilistic information content (PIC), i.e. to minimize the Shannon entropy, of the approximated subjective probability measure drawn from any bba. Justification and comparisons of $DSmP(.)$ w.r.t. $BetP(.)$ and to other transformations can be found in details in [8, 15], Vol. 3, Chap. 3. $DSmP$ transformation is defined[12] by $DSmP_\epsilon(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$ by

$$DSmP_\epsilon(X) = \sum_{Y \in 2^\Theta} \frac{\sum_{\substack{Z \subseteq X \cap Y \\ \mathcal{C}(Z)=1}} m(Z) + \epsilon \cdot \mathcal{C}(X \cap Y)}{\sum_{\substack{Z \subseteq Y \\ \mathcal{C}(Z)=1}} m(Z) + \epsilon \cdot \mathcal{C}(Y)} m(Y) \quad (17)$$

where $\mathcal{C}(X \cap Y)$ and $\mathcal{C}(Y)$ denote the cardinals of the sets $X \cap Y$ and Y respectively; $\epsilon \geq 0$ is a small number which allows to reach a highest PIC value of the approximation of $m(.)$ into a subjective probability measure. Usually $\epsilon = 0$, but in some particular degenerate cases, when the $DSmP_{\epsilon=0}(.)$ values cannot be derived, the $DSmP_{\epsilon>0}$ values can however always be derived by choosing ϵ as a very small positive number, say $\epsilon = 1/1000$ for example in order to be as close as we want to the highest value of the PIC. The smaller ϵ, the better/bigger PIC value one gets. When $\epsilon = 1$ and when the masses of all elements Z having $\mathcal{C}(Z) = 1$ are zero, $DSmP_{\epsilon=1}(.) = BetP(.)$.

[12] Here we work on classical power-set, but DSmP can be defined also for working with other fusion spaces, hyper-power sets or super-power sets if necessary.

4.3 Sketch of PCR-BIMM

We briefly summarize the five steps of our PCR-BIMM filter.

• **Step 0** (Initialization at $k = 0$): Same as Step 0 of BIMM.

• **Step 1** (Interaction-mixing): Same as Step 1 of BIMM except that the predicted bba $\mathbf{m}_k^-(.)$ is computed by (15) instead of (11), that is

$$\mathbf{m}_k^-(.) \triangleq \mathbf{M}_t' \cdot \mathbf{m}_{k-1}(.) \quad (18)$$

and the derivation of the mixing probability $\mu_{i|j}(k-1|k-1) = P\{M_i(k-1)|M_j(k), \mathbf{Z}^{k-1}\}$ of classical IMM is replaced by the DSmP probability drawn from $\mathbf{m}_{k-1|k-1}(.|M_j(k))$, that is:

$$\mu_{i|j}(k-1|k-1) = DSmP_\epsilon(M_i(k-1)|M_j(k), \mathbf{Z}^{k-1})$$

where $DSmP_\epsilon(.)$ is calculated with the transformation (17) using $\mathbf{m}_{k-1|k-1}(.|M_j(k))$ given by (12).

• **Step 2**: Same as IMM Step 2.

• **Step 3** (Mode bba update): The updated bba $\mathbf{m}_k(.)$ of modes is computed from the PCR5 (or eventually PCR6) rule, denoted $\oplus$, of the predicted bba $\mathbf{m}_{k-1}^-(.)$ given in (15) with bba's $\mathbf{m}_{k,j}(.)$, $j = 1, 2, \ldots r$ by

$$\mathbf{m}_k(.) = [\mathbf{m}_{k,1} \oplus \ldots \oplus \mathbf{m}_{k,r} \oplus \mathbf{m}_{k-1}^-](.) \quad (19)$$

where the *observed bba's* $\mathbf{m}_{k,j}(.)$ for $j = 1, \ldots, r$ are given as in BIMM by (14).

• **Step 4** (Global estimation for output purpose): The global estimate $\hat{\mathbf{x}}(k|k)$ and the covariance of estimation error $\mathbf{P}(k|k)$ are given as in step 4 of classical IMM by taking $\mu_j(k) = DSmP_\epsilon\{M_j(k)|\mathbf{Z}^k\}$ computed from the updated bba $\mathbf{m}_k(.)$ by (17).

Remark: This preliminary version of PCR-BIMM is perfectible because it still shares several points with BIMM[13]. In particular, the Step 3 of PCR-BIMM calculates, as in BIMM, $\mathbf{m}_{k,j}(.)$ with a model based on likelihoods $\Lambda_j(k)$ whose strong justification is missing. Further investigations will be done to improve this step 3, as well as the Step 1 to get better performances of PCR-BIMM (if possible) in a future research.

5 Simulation results

In this section, we present the application of the PCR-BIMM to a ground target tracking problem. We consider a vehicule localized in $(1000m, 5000m)$ in the cartesian referential (X, Y). We simulate a ground sensor located in (0,0) which is able to detect the moving target in range ρ and azimut θ. The gaussian measurement noise is supposed to be white and centered with the covariances $\sigma_\rho = 20\,m$ and $\sigma_\theta = 0.008\,rad$. The sampling time is fixed to 2 seconds. For tracking the ground target we only consider two motion models.

[13] In particular, the GBT is still used in Step 2 of PCR-BIMM.

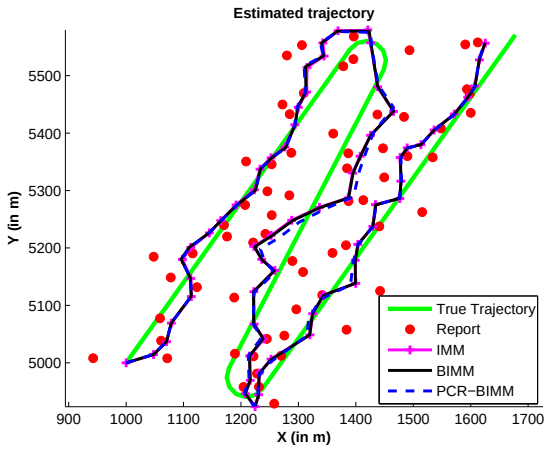


Figure 1: True target trajectory and estimated trajectories.

A constant velocity motion model called CV 1, with a small noise $\sigma_{CV_1} = 1\,m.s^{-2}$ and another constant velocity motion model called CV 2, with a bigger noise $\sigma_{CV_2} = 4\,m.s^{-2}$ to palliate the target maneuver. The initial state for each IMM, BIMM[14] and PCR-BIMM is the true initial target state $\mathbf{x}(0)$. The transition Matrix $\mathbf{P}_t$ is equal to :

$$\mathbf{P}_t = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix} \tag{20}$$

and the mass transition matrix $\mathbf{M}_t$ for the BIMM and PCR-BIMM is same as in the paper [13]. The initial motion model mass is represented by the vacuous mass function.

To compare the performances between the algorithms we used the root mean square error (RMSE) in location and velocity (figure 2) and the mean of the motion models probability obtained with 100 Monte-Carlo runs (figures 3, 4, 5). The first remark is, there is no significant improvement by using the belief function in the IMM. In fact, the RMSE of the IMM, BIMM and PCR-BIMM are globally the same. However, we can observe a short difference of the PCR-BIMM error after the target maneuvers between the time intervals $[20, 30]$ and $[40, 50]$. This observation carries along the second remark: the motion model transition duration is longer with the IMM (figure 3) and BIMM (figure 4) than the PCR-BIMM (figure 5). Then with the taken parameters for this simulation, the PCR-BIMM appears to be a good and fast detector of the motion models transition. However, its computed motion models probability is inferior to the probability obtained with the IMM and BIMM. More investigations need to be done to see if it is possible (and how) to improve PCR-BIMM in order to preserve both the good performance of the maneuver detection and in the same time and get higher probability when the target is moving in the same mode.

[14] Our BIMM implementations uses algorithm described in section 3 with (15) and additional normalization step $m_k(.)$ in (7) since otherwise the BIMM algorithm doesn't work at all due to the problem mentioned in section 2.

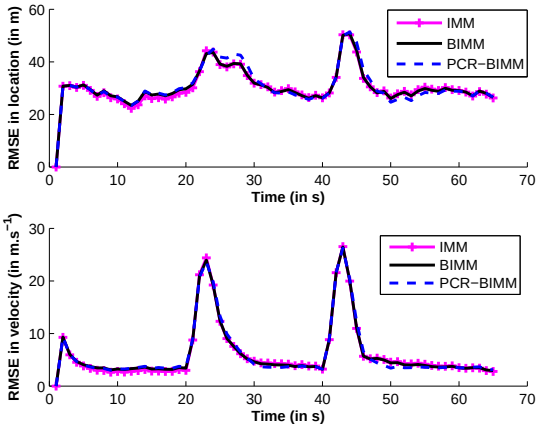


Figure 2: Root Mean Square Error.

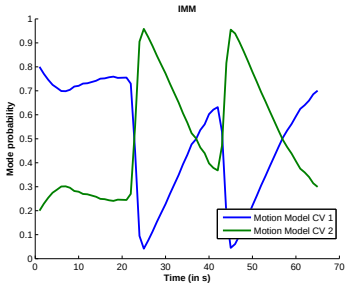


Figure 3: Motion Model Probability of the IMM.

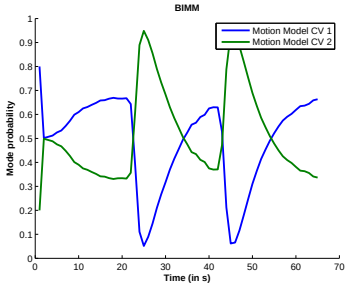


Figure 4: Motion Model Probability of the BIMM.

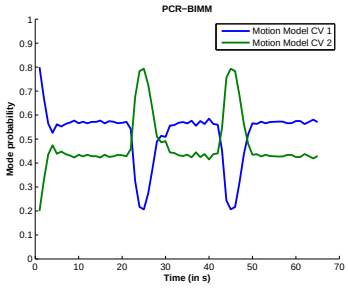


Figure 5: Motion Model Probability of the PCR-BIMM

6 Conclusions

In this paper, we have examined in details the recent BIMM algorithm and have corrected a mistake in it, and also identified some of its limitations. To palliate the problems of BIMM algorithm, we have developed a more efficient belief-based algorithm, called PCR-BIMM, based on the Proportional Conflict Redistribution fusion rule and on the DSmP probabilistic transformation to replace the conjunctive rule and the pignistic transformation used in BIMM. The derivation of the predicted bba of modes done incorrectly in BIMM is also fixed in our PCR-BIMM filter. The perfomances of PCR-BIMM with respect to the (corrected) BIMM and to the classical IMM have been evaluated from a simple maneuvering target tracking scenario through Monte-Carlo simulations. The results obtained in this paper show the ability of the PCR-BIMM to track maneuvering targets and also to improve the maneuver detection. It is important to note that such PCR-BIMM filter can be considered as more robust than IMM since PCR-BIMM requires less specific prior information than IMM. Nevertheless, PCR-BIMM provides globally the same RMS estimation errors performances as those obtained with the classical IMM which requires more specific prior information. Application of PCR-BIMM for tracking multiple maneuvering ground targets in a battlefield surveillance context is under investigation and results will be published in forthcoming papers.

References

[1] A. Appriou, *Classification based on multi-sensor uncertain data-fusion*, Agardograph 337 on Multisensor Multitarget Data Fusion, Tracking and Identification Techniques for Guidance and Control Applications, (in french), 1996.

[2] Y. Bar-Shalom, K. C. Chang, H. A. P. Blom, *Tracking a maneuvering target using input estimation vs. the interacting multiple model algorithm*, IEEE Trans on AES, Vol. 25, pp. 296–300, March 1989.

[3] Y. Bar-Shalom, X.-R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Artech House, Boston, MA, 1993.

[4] Y. Bar-Shalom, X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS Publishing, 1995.

[5] H. A. P. Blom, *An efficient filter for abruptly changig systems*, Proc. of the 23rd Conf. on Decision and Control, Las vegas, NV, USA, Dec., 1984.

[6] H. A. P. Blom, Y. Bar-Shalom, *The Interacting multiple model algorithm for systems with markovian switching coefficients*, IEEE Trans on AC, Vol. 33, No. 8, pp. 780–783, Aug. 1988.

[7] J. Dezert, A. Tchamova, F. Smarandache, P. Konstantinova, *Target Type Tracking with PCR5 and Dempster's Rules: A Comparative Analysis*, Proc. of Fusion 2006 Int. Conf., Firenze, Italy, July 2006.

[8] J. Dezert, F. Smarandache, *A new probabilistic transformation of belief mass assignment*, in Proceedings of Fusion 2008 Conference, Cologne, Germany, July 2008.

[9] J. Dezert, F. Smarandache, *Non Bayesian conditioning and deconditioning*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[10] X. R. Li, Y. Bar-Shalom, *Multiple-model estimation with variable structure*, IEEE Trans. on AC, Vol. 41, No. 4, pp. 478–493, April 1996.

[11] X. R. Li, *Engineer's guide to variable-structure multiple-model estimation for tracking*, in Y. Bar-Shalom and W. D. Blair, eds., Multitarget-Multisensor Tracking: Applications and Advances, Vol. 3, Chapter 10, Artech House, Boston, pp. 499–567, 2000.

[12] X. R. Li, V. P. Jilkov, *Survey of maneuvering target tracking. Part V: multiple-model methods*, IEEE Trans. on AES, Vol. 41, No. 4, pp. 1255–1321, Oct. 2005.

[13] G. Nassreddine, F. Abdallah, T. Denœux, *A state estimation method for multiple model systems using belief function theory*, Proc. of Information Fusion conf. (FUSION 09), Seattle, Washington, USA, July 2009.

[14] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, 1976.

[15] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3, American Research Press, 2004–2009. http://fs.gallup.unm.edu/DSmT.htm

[16] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, (submitted to Fusion 2010 conference).

[17] Ph. Smets, *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in AI, Vol. 5, pp. 29–39, 1990.

[18] Ph. Smets, *Belief Functions: The Disjunctive Rule of Combination and the Generalized Bayesian Theorem*, Int. Journal of Approximate Reasoning, Vol. 9, pp. 1–35, August 1993.

Implementation of Approximations of Belief Functions for Fusion of ESM Reports within the DSm Framework

Pascal Djiknavorian
Pierre Valin
Dominic Grenier

Originally published as Djiknavorian P., Valin P., Grenier D., *Implementation of Approximations of Belief Functions for Fusion of ESM Reports within the DSm Framework*, in Proc. of Fusion 2010, Edinburgh, Scotland, UK, 26-29 July 2010, and reprinted with permission.

Abstract - Electronic Support Measures consist of passive receivers which can identify emitters which, in turn, can be related to platforms that belong to 3 classes: Friend, Neutral, or Hostile. Decision makers prefer results presented in STANAG 1241 allegiance form, which adds 2 new classes: Assumed Friend, and Suspect. Dezert-Smarandache (DSm) theory is particularly suited to this problem, since it allows for intersections between the original 3 classes. However, as we know, the DSm hybrid combination rule is highly complex to execute and requires high amounts of resources. We have applied and studied a Matlab implementation of Tessem's k-l-x, Lowrance's Summarization and Simard's approximation techniques in the DSm theory for the fusion of ESM reports. Results are presented showing that we can improve on the time of execution while maintaining or getting better rates of good decisions in some cases.

Keywords: Dezert-Smarandache Theory, ESM, approximations, Belief functions.

1 Introduction

In terms of classification, the Dezert-Smarandache theory (DSmT) can become quite useful, especially for the direct resolution of classification for cases of hierarchical classes structures. For instance, we have the case of the allegiance classification structure suggested by STANAG 1241 where a structure of five classes (3 main classes and 2 derived classes) is required. The DSmT is able to output to any of those classes without modifications to its fusion process.

However, this example is still a simple one and both DSmT theories, with or without approximation, can solve it quite easily, which wouldn't be the case for classification problems of higher dimension. By dimension we mean the cardinal of the frame of discernment. In fact, the DSmT can become highly complex and computationally prohibitive as soon as we reach a dimension of 6. That is a classification of a problem having six main classes and up to, in the worst case scenario, a total of 7,828,353 possible derived classes.

Various avenues of research have been tried to avoid or address this complexity problem [10, 13, 18]. However, even just counting the number of possible classes is still an active problem in mathematics known as the Dedekind problem, or the problem of counting antichains [9, 18].

In this paper, we study the use of an approximation technique to restrain the staggering amount of data that the DSmT can generate in its fusion process. More specifically we have chosen Tessem's klx approximation technique [4], Lowrance's Summarization [19], Simard's and al technique [3, 7, 8] and used them into the DSmT with the DSm hybrid combination rule (DSmH). We have also experimented with the fusion process while using the approximation technique and compared it to the case without an approximation technique to analyze how it affects the quality of the decision process. More specifically, we will compare the good decision rate in the two cases, with and without the use of approximation.

1.1 Realistic Case Study

Electronic Support Measures (ESM) consist of passive receivers which can identify emitters coming from a small bearing angle, which, in turn, can be related to platforms that belong to 3 classes: either Friend (F), Neutral (N), or Hostile (H). Decision makers prefer results presented in STANAG 1241 allegiance form, which adds 2 classes: Assumed Friend (AF), and Suspect (S).

The DSm theory is particularly suited to this problem, since it allows for intersections between the original 3 classes of allegiance. In this way an intersection of Friend and Neutral can lead to an Assumed Friend, and an intersection of Hostile and Neutral can lead to a Suspect. This structure of allegiances will be referred to as STANAG allegiance [11].

Figure 1 displays a visual representation of a possible interpretation of STANAG allegiance in DSmT. We can see that even though the input consists only of three classes, we are able to give an output into five classes. For example, here we have the class 'Suspect', which could be the result obtained after fusing 'Hostile' with 'Neutral'. We also have the class 'Assumed Friend', which could be the result obtained after fusing 'Friend' with 'Neutral'. Note that this case example has the intersection $F \cap H = \emptyset$, the null set, which is a constraint in DSm, leading to the use of its hybrid rule. This case example would be relevant for peace-keeping missions where Hostile and Friendly forces aren't

likely to be close one to another. We will be working on that case, with F∩H = Ø.

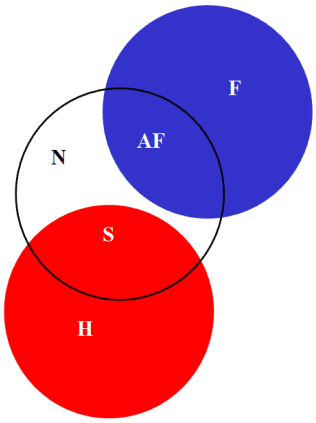


Figure 1. Venn diagram for the STANAG allegiances.

2 Dezert-Smarandache Theory

The DSm theory uses the language of masses assigned to each declaration from a sensor (in our case, the ESM sensor). In DSm theory, all unions and intersections are allowed for a declaration. For our case of cardinality 3, $\Theta = \{\theta_1, \theta_2, \theta_3\}$, with $|\Theta| = 3$, D^Θ is still of manageable size, namely has a cardinality of 19 [10]. In DSm theory, a constraint like the one that was imposed by Figure 1, namely that $F \cap H \equiv \theta_1 \cap \theta_3 = \emptyset$ is treated by the DSm hybrid combination rule (DSmH) below:

$$m(A) = \phi(A)\,[\,S_1(A) + S_2(A) + S_3(A)\,] \qquad (1)$$

The reader is referred to a series of books [10, 13, 17] on DSm theory for lengthy descriptions of the meaning of this formula (note that the function ϕ is not to be confused with the empty set). A three-step approach was proposed in [12], which is used here. The incoming sensor reports are either: Friend ($F = \theta_1$), Neutral ($N = \theta_3$) or Hostile ($H = \theta_3$), Figure 1 has the interpretation of the five classes:

$$\text{Friend} = \{\theta_1 - \theta_1 \cap \theta_2\} \qquad (2)$$
$$\text{Hostile} = \{\theta_3 - \theta_3 \cap \theta_2\} \qquad (3)$$
$$\text{Assumed Friend} = \{\theta_1 \cap \theta_2\} \qquad (4)$$
$$\text{Suspect} = \{\theta_2 \cap \theta_3\} \qquad (5)$$
$$\text{Neutral} = \{\theta_2 - \theta_1 \cap \theta_2 - \theta_3 \cap \theta_2\} \qquad (6)$$

As in [15], we call STANAG-probability the pignistic probability assigned to the five classes shown by equations (2) to (6). We use the general pignistic transform, as shown by [10] or equation (7), to obtain the probability values of the sets used in those equations.

$$P\{A\} = \sum_{X \in D^\Theta} \frac{C_\mathcal{M}(X \cap A)\, m(X)}{C_\mathcal{M}(X)} \qquad (7)$$

Where $C_M(A)$, is the DSm cardinal of a set A. It accounts for the total number of partitions. Each of these partitions possesses a numeric weight equal to one. That weight, identical for each part makes them all equal. The DSm cardinal is used in the generalized pignistic transformation equation to redistribute the masse of a set A among all its partitions B such that B is included or equal to A.

3 Approximation technique

3.1 K-l-x approximation

The k-l-x approximation technique developed by Tessem [4] is designed to approximate Basic Probability Assignment (BPA) or mass function in Dempster-Shafer Theory (DST). Since DSm theory works directly with BPAs, applying the k-l-x approximation technique to the DSmH is quite straightforward and can be done without any changes.

This algorithm for approximation of BPAs involves three parameters: k the minimum number of focal elements to be kept, l the maximum number of focal elements to be kept and x the maximum threshold on the sum of the lost masses. It can be summarized as follows:

1. Select the k focal elements with highest masses;
2. While the sum of their masses is less than 1-x, and while their number is less than l, add the next focal element with highest mass.

3.2 Simard's and al. approximation

This truncation scheme [3, 7, 8] has had many minor variations over time. Similarly to k-l-x approximation, it was conceived to approximate BPA or mass function in DST. And as in k-l-x, we were able to transfered it to the DSm framework. Variants exist but all focus on preferentially keeping fused propositions with the smallest lengths (lowest cardinality) after passing 2 thresholding steps. The rule therefore involves 3 parameters: BPAmax, BPAmin and Nmax. It retains fused propositions according to the following rules:

1. All fused propositions with BPA > BPA_{max} are kept (thresholding step 1)
2. All fused propositions with BPA < BPA_{min} are discarded (thresholding step 2)

3. If the number of retained propositions in step 1 is smaller than N_{max}, retain by decreasing BPA, propositions of length 1, then if the number of retained propositions is smaller than N_{max},, retain by decreasing BPA, propositions of length 2, and so on for 3…
4. If the number of retained propositions is still smaller than N_{max}, retain propositions by decreasing BPA regardless of length.

3.3 Lowrance's approximation

Similar to the k-l-x procedure, the summarization method [19] (inspired by the summarization operation described in Bauer's research [5]) leaves the best-valued focal elements of the mass function under consideration unchanged. The numerical values of the remaining focal elements are accumulated and assigned to the set-theoretic union of the corresponding subsets of Θ. Here again, the technique was conceived to approximate BPA or mass function in DST, and we were able to transfer it to the DSm framework.

3.4 Implementation of approximations

The information coming from the sensor is a simple belief function giving a mass to an allegiance and the remaining mass to ignorance. The combination itself combines two belief functions, one is the information from the sensor at time *t*, the other contains past information within combination result from time *t-1*. The fusion process is realized dynamically. Since the information to combine from the sensor is a simple belief function the approximation is applied on the result of the combination.

4 A typical simulation scenario

The pre-requisites that a typical scenario must address are: (1) to be able to adequately represent the known ground truth, (2) to contain sufficient countermeasures (or miss-associations) to be realistic and to test the robustness of the theories, (3) to only provide partial knowledge about the ESM sensor declaration, which therefore contains uncertainty, (4) to be able to show stability under countermeasures, yet (5) to be able to switch allegiance when the ground truth does so.

The following scenario parameters have therefore been chosen accordingly: (1) ground truth is FRIEND for the first 50 iterations of the scenario and HOSTILE for the last 50, (2) the number of correct associations is 80%, corresponding to countermeasures appearing 20% of the time, in a randomly selected sequence, (3) the ESM declaration has a mass (confidence value in Bayesian terms) of 0.8, with the rest of the mass being assigned to the ignorance (the full set of elements, namely Θ).

This scenario will be the one addressed in the next section, while a Monte-Carlo study is described in the subsequent sections. Each Monte-Carlo run corresponds to a different realization using the above scenario parameters, but with a different random seed. The chosen scenario is depicted in Figure 2.

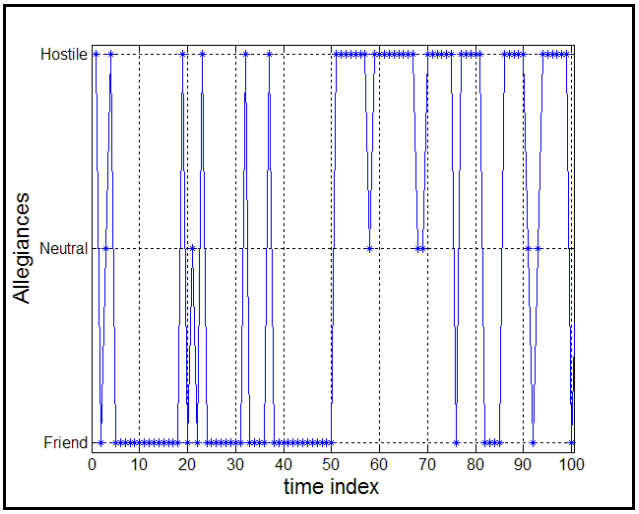


Figure 2. Chosen scenario.

Roughly 80% of the time the ESM declares the correct allegiance according to ground truth, and the remaining 20% is roughly equally split between the other two allegiances. Note that these percentages of occurrences are from a statistical point of view only, so that in the long run a large amount of randomly generated scenarios would amount to these ratios. There is an allegiance switch at the 50th time index, and the selected randomly selected seed in the above generated scenario generates a rather unusual sequence of 4 false Friend declarations starting at time index 82 (when actually Hostile is the ground truth).

4.1 Results for the simulated scenario

Before presenting the results, it should be noted that the original form of the DSmH tends to accumulates masses to intersections as is the case for any rule based on conjunction [14]. An ad hoc solution exists [3, 7, 8], and consists in renormalizing after each fusion step by giving a value to the complete ignorance which can never be below a certain factor (chosen here to be 0.04 as research in [14] shows that this value is appropriate for this case while being high enough to avoid the accumulation but still low enough not to interfere with the combination's performances). That solution was originally developed to the well-known problem of DST combination, which tends to be overly optimistic, which in turn prevents it to react quickly to changes of allegiances. For more on the behavior of the DSmH on similar cases the reader is referred to [14, 15, 16], as we are focused on exploring the effect of approximations on DSm here.

Since the whole idea behind using DSm was to present the results to the decision maker in the STANAG allegiance format, the result of Figure 3 would be used. For the DSmH

[10], it was suggested to use the Generalized Pignistic Probability, which is based on the pignistic transformation [6, 10], in order to make a decision on a singleton belonging to the input ESM-allegiance.

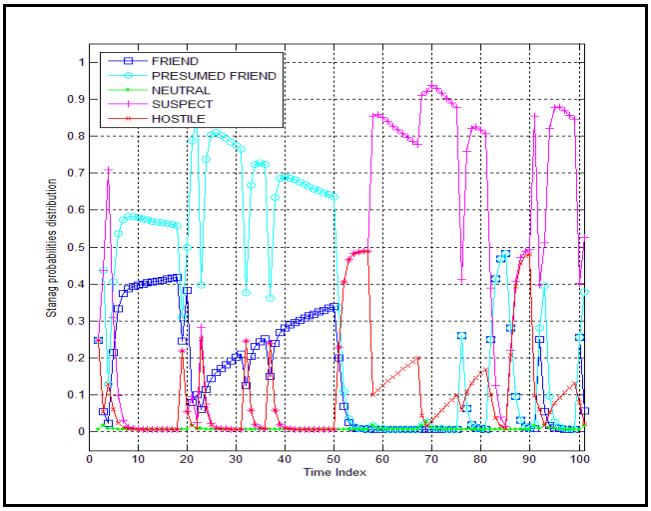


Figure 3. DSmH result for the chosen scenario.

The decision maker would clearly be informed that miss-associations have occurred, since Assumed Friend dominates for the first 50 time indices and Suspect for the latter 50. The Friend declarations starting at time index 82 cause confusion, as it should. The change in allegiance at time index 50 is detected quickly. What is even more important is that F and AF are clearly preferred for the first 50 time indexes and S and H for the last 50, as they should.

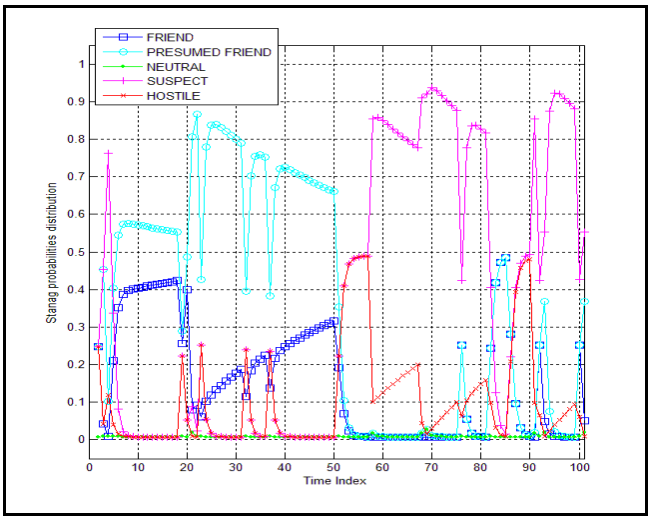


Figure 4. Approximated DSmH result for the same scenario with k-l-x = (5, 6, 0.2)

We can gather from Figure 4 and Figure 5 that the DSmH and the approximated DSmH have very similar behaviors. In fact, one has to look at the figures very closely to perceive the differences. We can see that in the first half of the approximated version, the assumed friend allegiance is slightly favored to the friend allegiance. Near the end of the scenario the hostile allegiance is favored to the suspect allegiance. However, in both cases, even if the smallness of the change could possibly affect our decision, the STANAG-probability still seems to stay within the same type of allegiance in the sense that a friend and a target of assumed friend allegiance would both inspire a friendly response on our part. The same can be said for a target of suspect or hostile allegiance that would both inspire a hostile or defensive response on our part. In short, we can easily proceed with the approximation and still be able to make the same decision the same way.

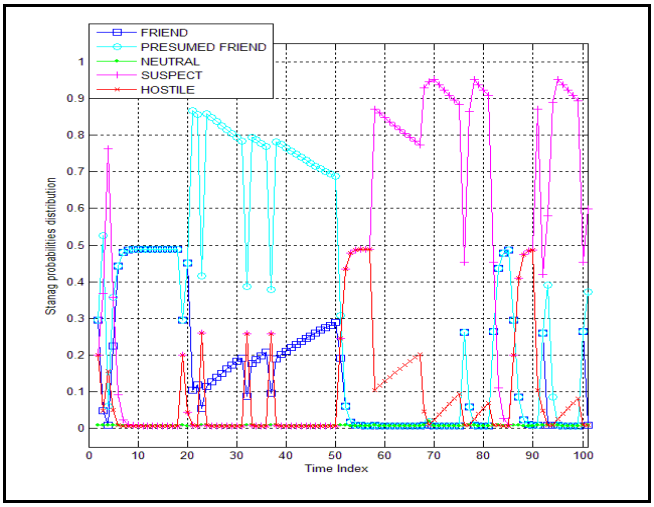


Figure 5. Approximated DSmH result for the same scenario with k-l-x = (3, 6, 0.2)

4.2 Effects of varying the k-l-x parameters

We've realized the scenario for various values of k-l-x for k ∈ [3, 10], l ∈ [6, 12] and x ∈ [0.2, 0.4]. For the cases where we had k=8, no changes in l and x had impact, and compared to the DSmH, we've only noticed a very small variation at the start and end of the simulated scenario. For the cases where we had k=6, no changes in l and x had impact and compared to DSmH, there was only very little variation in value throughout the scenario. The same is true for the cases with k=5, with the Figure 4 showing the results for that case. The amplitude of the variation between DSmH and the approximated version continues to increase as the k value diminishes.

We finally begin to notice small changes with x=0.2 as opposed to 0.3 or 0.4 when we reach k=4. However, the impact of having x at 0.2 is small and contained at the start of the scenario, where it gives more weight to the suspect class at the expense of the hostile class. For the cases with k=3, the impact of the change on x going to 0.2 was more significant and lasted throughout most of the scenario's duration. Also, while for cases of k ∈ [4, 8] the behavior of the curves were all very similar one to another, when we reach k=3, we observe a partial loss of smoothness, hence a more reactive behavior toward countermeasures and allegiance change. Figure 5 shows the case of the simulated

scenario for an approximated DSmH with klx = (3, 6, 0.2). Note that in all our experimentations for our chosen scenario the l parameter never had any visible impact.

5 Monte-Carlo Simulations with k-l-x approximation

Although a special case such as the one described in the previous section offers valuable insight, one might question if the conclusions from that one scenario pass the test of multiple Monte-Carlo scenarios. This question is answered in this section.

In order to expend the parameter space, we have realized the simulations of the current section to 80 and 90% for the ESM certainty, and with an ESM confidence at 80% and an ignorance threshold at 0.04 as before. The number of Monte-Carlo runs was set to 100. The randomly generated ESM stream of reports used for both the DSmH and the approximated DSmH are all the same so that we can freely compare the effects of the use of the approximation, and the impact of the variation of its parameters.

As for the choice of a the graphical display to highlight the results of our simulations, we went with the rate of good decisions, where a good decision is as we have mentioned earlier, when we conclude to be friendly toward a friendly behaving target, when the ground truth is of class friend. A friendly-behaving target is a target that is concluded to be a friend or an assumed friend. We also have a good decision when we conclude to be hostile toward a hostile behaving target, when the ground truth is of class hostile. A hostile-behaving target is a target that is concluded to be a hostile or a suspect. A decision is made by taking the set of maximum STANAG-probability.

5.1 Effects of varying the k-l-x parameters

Simulations were done on a computer with a Phenom II 955 processor with 8 GB of memory. We should keep in mind that it is the relative time of execution which is important here. For figures 7 to 11, the simulations had a value of 80% for the ESM certainty and the value of the x parameter was maintained at 0.2 since changing it had no impact on good decision rate.

Figure 7 and Figure 8 show us the effect of the approximation from the good decision rate point of view when compared with the DSmH case from Figure 6. Like for the typical simulated scenario from previous section, l had no visible impact, and x had a limited impact only as the k parameter went below 4. As for the k parameter, it started having an impact when we reached 6, where the impact was on only three iterations. As the k parameter reaches 5, a very slight positive impact throughout the whole simulation can be seen. As for k=4 and k=3, we have a slight deterioration of the good decision rate but it is still very small and rather insignificant considering the gain in time execution as Figure 10 shows us. For the cases with an ESM confidence at 90%, all the approximated results, have no significant impact on the good decision rate, except with klx = (3, 8, 0.2) where we had minimal impact.

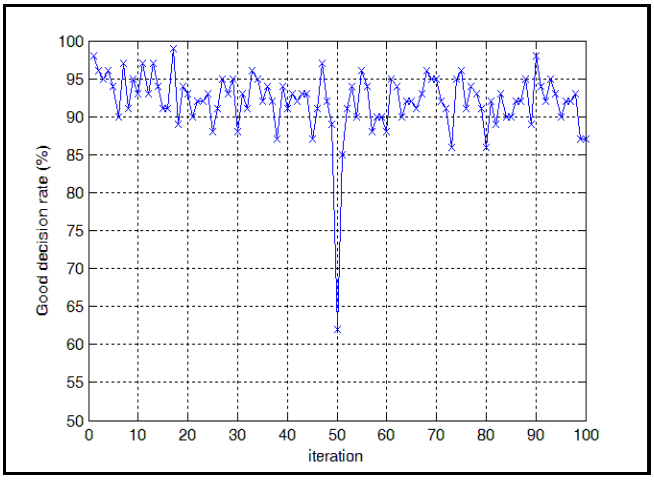


Figure 6. DSmH result after 100 Monte-Carlo runs.

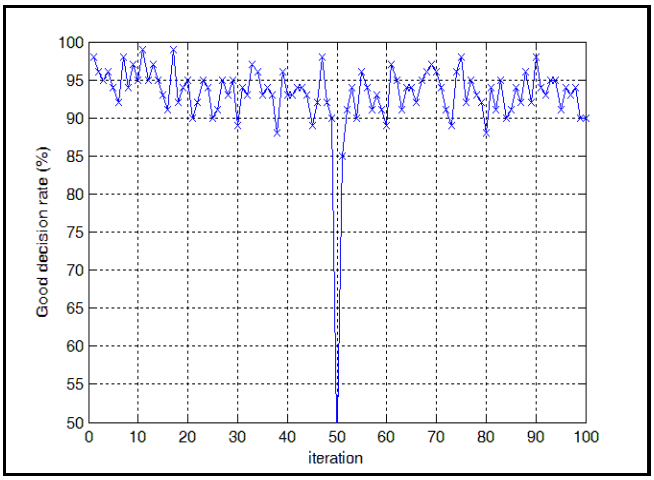


Figure 7. Approximated DSmH result with k-l-x = (5, 8, 0.2) for the same Monte-Carlo simulation.

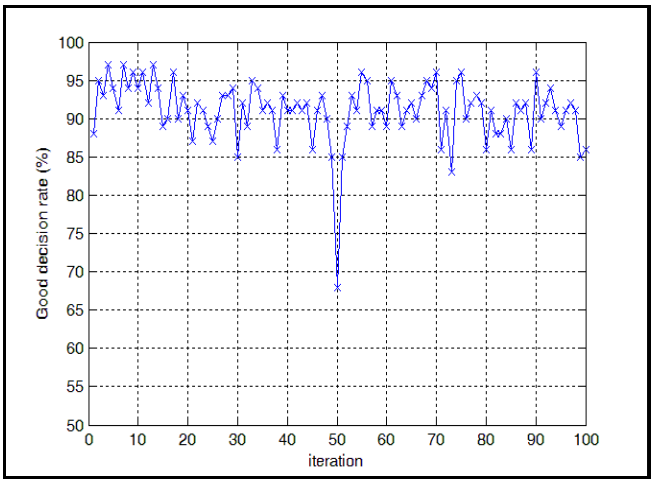


Figure 8. Approximated DSmH result with k-l-x=(3,8,0.2) for the same Monte-Carlo simulation.

We have the time of execution versus k and l parameters from the klx approximation technique on Figure 9 and Figure 10. Specifically, Figure 9 has the curve of the time of execution of the combination and approximation process only. The x-y plane, valued at 325.97 seconds on Figure 9 indicates the time from which the approximation process provides a higher gain in time than the time it consumes. It is the time of execution of the DSmH without approximation.

We can see that the k parameter has to reach 5 before we start seeing an improvement. Before that value, the approximation takes more time to execute than it helps us gain. We can achieve a 30% improvement on time of execution when we reach k = 3. The parameter l has no impact on time. The absence of impact of the l parameter is suspected to be caused by the fact that this simulated scenario case uses simple support functions as inputs.

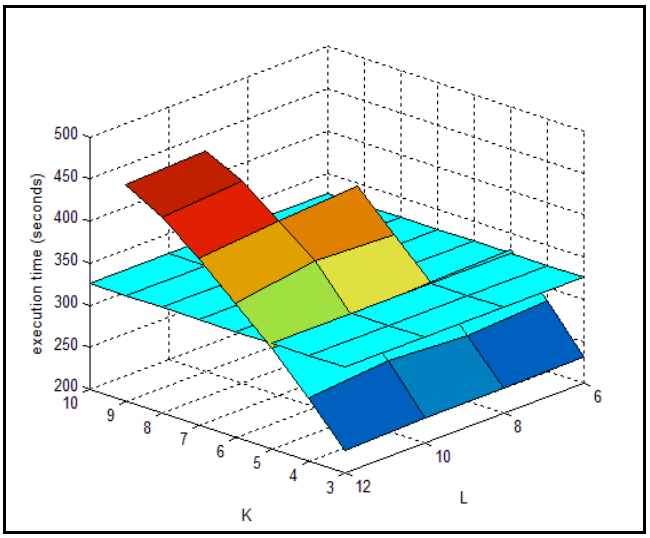


Figure 9. Execution time for the combination and approximation processes.

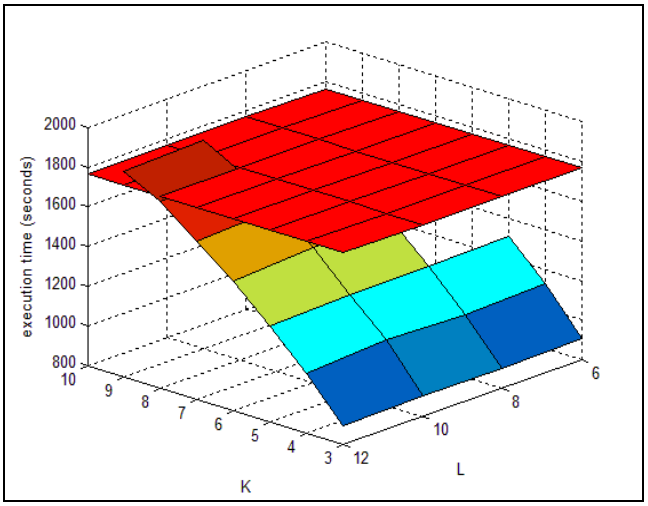


Figure 10. Execution time for the whole simulation.

In Figure 10, we have the curve of the time of execution for the whole simulation which, on top of the combination and approximation processes, includes the generalized pignistic transformation (GPT) which is used in the decision process. Above 95% of the extra time of execution, when compared to figure 10, is composed of the GPT.

In Figure 10, the x-y plane, representing the time of execution of the simulation without approximation, is valued at 1767.6 seconds. We can see that we can have a 50% reduction in time of execution when we reach k=3 and that l has no impact. As we compare Figure 9 and Figure 10, we see that the GPT is the step that benefits the most from the approximation process.

6 Monte-Carlo simulations using various approximation rules

In order to expend the analysis furthermore, we have realized the simulations of the current section with Monte-Carlo runs set to 1000. Also, we've expended the analysis to Simard's summarization, and Simard's truncation techniques with the same stream of reports to fuse. Hence, both the DSmH and the approximated DSmH will have the same dataset so that we can freely compare the effects of the use of the approximation, and the impact of the variation of its parameters.

Figure 11, which shows results using Lowrance's approximation technique lets us see the inability of the technique to get better good decision rates than the non approximated combination. The following figures shows that k-l-x, and Simard's Truncation are both able to get, depending on the chosen parameters, better results of good decision rates, than the scenario without approximation.

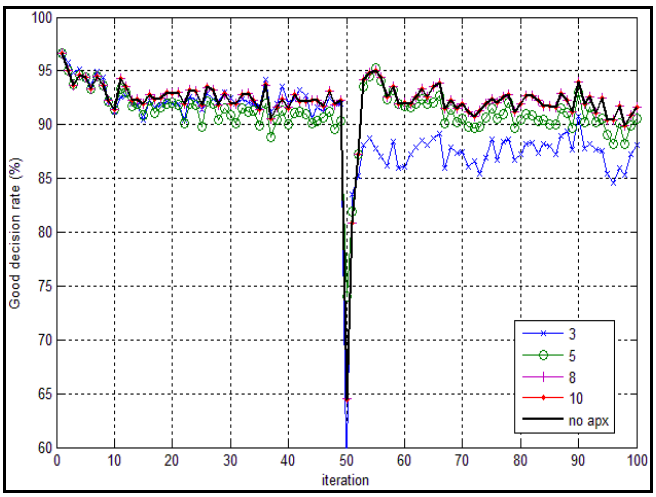


Figure 11. DSmH using Lowrance's apx. (3/5/8/10).

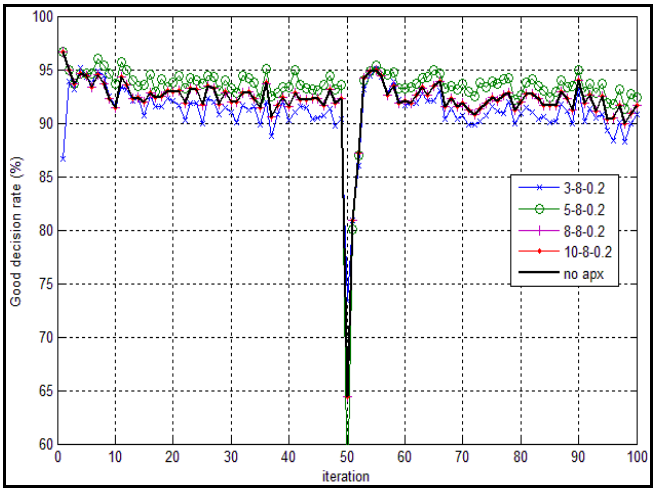


Figure 12. DSmH using k-l-x apx. (3/5/8/10–8–0.2).

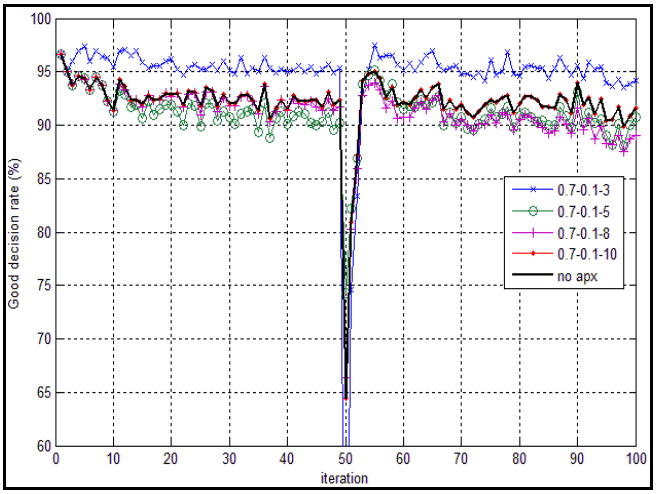


Figure 15. DSmH using Simard's apx. (0.7–0.1–3/5/8/10).

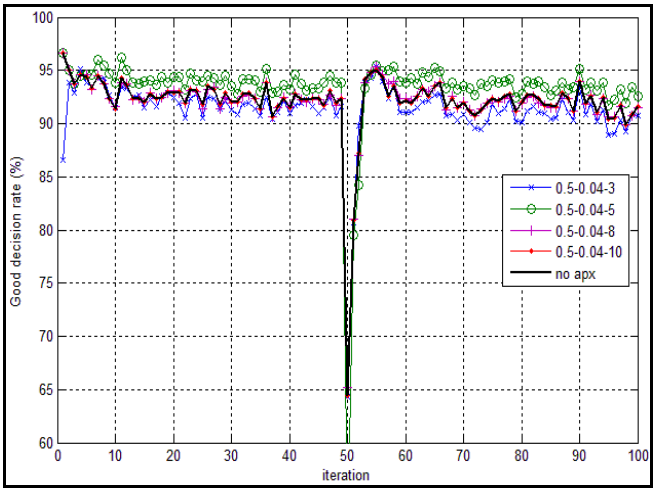


Figure 13. DSmH using Simard's apx. (0.5-0.04-3/5/8/10).

About the mean time of execution of the combination and approximation step for realistic scenario, we have found that for a parameter 'K' below, or equal, to 5, we were able to execute faster than without approximation. And when looking at previous figures, we see that, too low (K~3), the approximation isn't as good as without approximation, and that at a value of 5, we were always at higher good decision rates than the case without approximation.

So not only we have found a case executing faster than without approximation, but we've also found ourselves a case where it performs better in terms of good decision rate. That is for approximation techniques different from Lowrance's, and limited, until proven differently, to this case, and for DSmH.

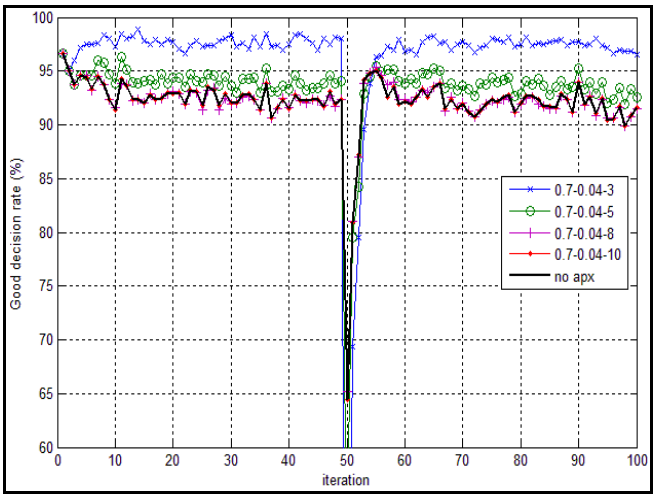


Figure 14. DSmH using Simard's apx. (0.7–0.04–3/5/8/10).

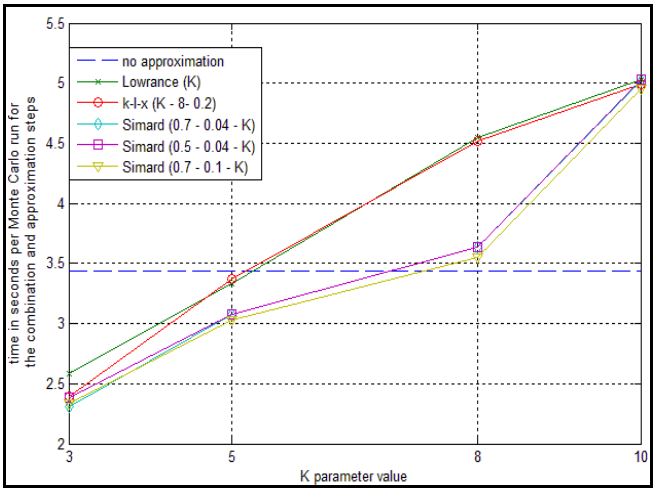

time in seconds per Monte Carlo run for the combination and approximation steps
K parameter value
no approximation
Lowrance (K)
k-l-x (K - 8- 0.2)
Simard (0.7 - 0.04 - K)
Simard (0.5 - 0.04 - K)
Simard (0.7 - 0.1 - K)

Figure 16. Combination and approximation execution times in seconds per Monte-Carlo run.

7 Conclusions

The previous sections display the behavior for different cases of klx approximation on the same simulated ESM data (see Figure 6, Figure 7, and Figure 8). It also shows the time of execution of each of those simulations. From those results we can conclude that we can successfully attain the same good decision rate with DSmH as with an approximated DSmH for the chosen scenario, while achieving lower times of execution including the time to approximate when we reach a certain level of approximation. Those results are confirmed by the experimentation done on another simulated dataset lasting for 1000 Monte-Carlo runs. (see Figure 12)

We've also explored the behavior of Lowrance's summarization approximation and Simard's truncation methods while using the same dataset also on a thousand Monte-Carlo runs. From what we can see on Figure 11, the summarization is able, with the careful choice of its parameter, to reach good decision rate of the combination rule without approximation, however, it seems to be rarely able to do better and can do much worst. Simard's truncation method (see Figure 13, Figure 14 and Figure 15) on the other hand is able to get around 5% better good decision rates, depending on the choice of the approximation parameters. It can also get the same rates or a little less than the combination rule without approximation.

When considering results of time of execution as shown on Figure 16 we gather that, while being able to execute faster than the combination rule without approximation, we can get better decision rates. The 'K' parameter value of approximation of each rule, when at 3 or 5, gave us highest decision rates for Simard's truncation method or k-l-x approximation technique. Note that some times, parameter K had to be set at 3, other times at 5, depending on chosen technique and the other parameters, to reach highest decision rate.

Future work considered includes the exploration of the use of Bauer's D1 approximation [5] in DSmT. Even if it adds to the number of operations and in the complexity of the system, it would be interesting to see if the gain acquired by approximating is sufficient to counter the increase in complexity. We are also interested to see if it is able to give even better good decision rates than the other methods of approximation.

References

[1] A.P. Dempster,. *Upper and lower probabilities induced by a multivalued mapping*, Ann. Math. Statist. 38, 1967. pp. 325–339.

[2] G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[3] M.A. Simard, P. Valin and E. Shahbazian *Fusion of ESM, Radar, IFF and other Attribute Information for Target Identity Estimation and a Potential Application to the Canadian Patrol Frigate*, AGARD 66th Symposium on Challenge of Future EW System Design, 18-21 October 1993, Ankara (Turkey), AGARD-CP-546, pp. 14.1-14.18.

[4] B. Tessem, *Approximations for efficient computation in the theory of evidence*, Artificial Intelligence, vol. 61, pp. 315–329, June 1993.

[5] M. Bauer, *Approximation Algorithms and Decision Making in the Dempster-Shafer Theory of Evidence-An Empirical study*, International Journal of Approximate Reasoning, vol. 17, no. 2-3, pp. 217–237, 1997.

[6] Ph. Smets, *Data Fusion in the Transferable Belief Model*, Proceedings of the 3rd International Conference on Information Fusion, Fusion 2000, Paris, July 10-13, 2000, pp PS21-PS33.

[7] D. Boily, and P. Valin, *Truncated Dempster-Shafer Optimization and Benchmarking*, in Sensor Fusion: Architectures, Algorithms, and Applications IV, SPIE Aerosense 2000, Orlando, Florida, April 24-28 2000, Vol. 4051, pp. 237-246.

[8] D. Boily, and P. Valin, *Optimization and Benchmarking of Truncated Dempster-Shafer for Airborne Surveillance*, NATO Advanced Study Institute on Multisensor and Sensor Data Fusion, Pitlochry, Scotland, United Kingdom, June 25 – July 7 2000. Kluwer Academic Publishers, NATO Science Series, II. Mathematics Physics and Chemistry – Vol. 70, pp. 617-624.

[9] R. Fidytek, A.W. Mostowski, R. Somla and A. Szepietowski. *Algorithms counting monotone Boolean functions*, Information Processing Letters, Vol. 79, Issue 5, pp. 203-209, 15 September 2001.

[10] F. Smarandache, J. Dezert, editors. *Advances and Applications of DSmT for Information Fusion*, vol. 1, American Research Press, 2004.

[11] STANAG 1241 (2005). *NATO Standard Identity Description Structure for Tactical Use*, North Atlantic Treaty Organization, April 2005.

[12] P. Djiknavorian, and D. Grenier, *Reducing DSmT hybrid rule complexity through optimisation of the calculation algorithm*, in Advances and Applications of DSmT for Information Fusion, Collected Works edited by F. Smarandache, J. Dezert,, Volume 2, American Research Press, 2006.

[13] F. Smarandache, J. Dezert, editors. *Advances and Applications of DSmT for Information Fusion*, vol. 2, ARP, 2006.

[14] P. Djiknavorian, *Fusion d'informations dans un cadre de raisonnement de Dezert-Smarandache appliquée sur des rapports de capteurs ESM sous le STANAG 1241*, Mémoire de maîtrise, Université Laval, 2008.

[15] P. Djiknavorian, P. Valin, and D. Grenier, *Dezert-Smarandache theory applied to highly conflicting reports for identification and recognition – Illustrative example of ESM associations in dense environments*, DRDC Valcartier TR 2008- 537, 34 pages.

[16] P. Djiknavorian, P. Valin, and D. Grenier, *Fusion of ESM allegiance reports using DSmT*, in Advances and Applications of DSmT for Information Fusion, Collected Works edited by F. Smarandache, J. Dezert,, Volume 3, American Research Press, 2009.

[17] F. Smarandache, J. Dezert, editors. *Advances and Applications of DSmT for Information Fusion*, vol. 3, American Research Press, 2009.

[18] T. Carroll, J. Cooper and P. Tetali, *Counting Antichains and Linear Extensions in Generalizations of the Boolean Lattice*, August 30, 2009. http://www.math.sc.edu/~cooper/calegbl.pdf

[19] Lowrance, J. D., Garvey, T. D., and Strat, T. M., *A framework for evidential reasoning systems*, Proceedings of the 5th National Conference of the American Association for Artificial Intelligence, Philadelphia, 896-903, Aug. 1986.

Multiple Ground Target Tracking and Classification with DSmT

Benjamin Pannetier
Jean Dezert

Originally published as Pannetier B., Dezert J., *Multiple ground target tracking and classification with DSmT*, Informatik 2010, 5th German Workshop in Sensor Data Fusion (SDF 2010), Leipzig, Germany, pp. 886-891, Sept 30-Oct 1st, 2010, and reprinted with permission.

Abstract: Based on our previous work we propose to track multiple ground targets with GMTI (Ground Moving Target Indicator) sensors as well as with imagery sensors. The scope of this paper is to fuse the attribute type information given by heterogeneous sensors with DSmT (Dezert Smarandache Theory) and to introduce the type results in the tracking process to improve its performances.

1 Introduction

Data fusion for ground battlefield surveillance is more and more strategic in order to create the situational assessment or improve the precision of fire control system. For this, we develop new ground target tracking algorithms adapted to GMTI (Ground Moving Target Indicator) sensors. In fact, GMTI sensors are able to cover a large surveillance area during few hours or more if several sensors evolve on the same operational theatre. Several references exist for the MGT (Multiple Ground Tracking) with GMTI sensors [**?**, 8] whose fuse contextual informations with MTI reports. The main results are the improvement of the track precision and track continuity. Our algorithm [6] is built with several reflexions inspired of this literature. The proposed VS-IMMC (Variable Structure Interacting Multiple Models) filter is extended in a multiple target context and integrated in a SB-MHT (Structured Branching - Multiple Hypotheses Tracking).

One way to enhance data associations is to fused data obtained by several sensors. The most easily approach is to consider the centralized fusion between two or more GMTI sensors. Another way is to introduce heterogeneous sensors in the centralized architecture in order to improve the data associations (by using the reports in location and its classification attribute) and palliate the poor GMTI sensor classification. In our previous works [6], the classification information of the MTI segments and IMINT segments (IMagery INTelligence) has been introduced in the target tracking process. The idea was to maintain aside each target track a set of ID hypotheses. Their committed belief are revised in real time with the classifier decision through a very recent and efficient fusion rule called proportional conflict redistribution (PCR).

In this paper, in addition to the measurement location fusion, we illustrate on a complex scenario our approach to fuse MTI classification type with image classification type associated to each report.

2 Motion & observation models

2.1 Constrained motion model

The target state $\mathbf{x}(k)$ at the current time t_k is defined in a local horizontal plane (O, X, Y) of a Topographic Coordinate Frame denoted *TCF*. The target state on the road segment s is defined by $\mathbf{x}_s(k)$ where the target position $(x_s(k), y_s(k))$ belongs to the road segment s and the corresponding heading $(\dot{x}_s(k), \dot{y}_s(k))$ is in its direction. The event that the target is on road segment s is noted $e_s(k) = \{\mathbf{x}(k) \in s\}$. Given the event $e_s(k)$ and according to a motion model M_i, the estimation of the target state can be improved by considering the road segment s. The constrained motion model M_i^s is build in such a way that the predicted state is on the road segment s and the gaussian noise is defined under the road segment constaint [6]. After the state estimation obtained by a Kalman filter, the estimated state is then projected according to the road constraint $e_s(k)$. This process is detailed in [6].

2.2 GMTI measurement model

According to the NATO GMTI format [5], the MTI reports received at the fusion station are expressed in the WGS84 coordinates system. The MTI reports must be converted in the *TCF*. A MTI measurement z at the current time t_k is given in the *TCF*. Each MTI report is characterized both with the location and velocity information (range radial velocity) and also with the attribute information and its probability that it is correct. We denote C_{MTI} the frame of discernment on target ID based on MTI data. C_{MTI} is assumed to be constant over the time and consists in a finite set of exhaustive and exclusive elements representing the possible states of the target classification. In this paper, we consider only 3 elements in C_{MTI} defined as $C_{MTI} = \{$*Tracked vehicle*, *Wheeled vehicle*, *Rotary wing aircraft*$\}$.

We consider also the probabilities $P\{c(k)\}$ $(\forall c(k) \in C_{MTI})$ as input parameters of our tracking systems characterizing the global performances of the classifier. The vector of probabilities $[P(c_1)\ P(c_2)\ P(c_3)]$ represents the diagonal of the confusion matrix of the classification algorithm assumed to be used. Let $\mathbf{z}^{\star}_{MTI}(k)$ the extended MTI measurements including both kinematic part and attribute part expressed by the herein formula:

$$\mathbf{z}^{\star}_{MTI}(k) \triangleq \{\mathbf{z}_{MTI}(k), c(k), P\{c(k)\}\} \tag{1}$$

2.3 IMINT motion model

For the imagery intelligence (IMINT), we consider two sensor types : a video EO/IR sensor carried by a Unmanned Aerial Vehicle (UAV) and a EO sensor fixed on a Unattended Ground Sensor (UGS). We assume that the IMINT reports $\mathbf{z}_{video}(k)$ at the current time t_k are expressed in the reference frame (O, X, Y) and give a location information and type

target. We assume that the video information given by both sensor types are processed by their own ground stations and that the system provides the video reports of target detections with their classification attributes. For the last point, a human operator selects targets on a movie frame and is able to choose its attribute with a HMI (Human Machine Interface). Based on the military symbology called 2525C [3], we build the frame of discernment for an EO/IR source denoted C_{video}. Each video report is associated to the attribute information $c(k)(\forall c(k) \in C_{video})$ with its probability $P\{c(k)\}$ that it is correct. As C_{MTI}, C_{video} is assumed to be constant over the time and consists in a finite set of exhaustive and exclusive elements representing the possible states of the target classification.

Let $\mathbf{z}^{\star}_{video}(k)$ be the extended video measurements including both kinematic part and attribute part expressed by the following formula $(\forall c(k) \in C_{video})$:

$$\mathbf{z}^{\star}_{video}(k) \triangleq \{\mathbf{z}_{video}(k), c(k), P\{c(k)\}\} \tag{2}$$

The attribute type of the image sensors belongs to a different and better classification than the MTI sensors.

2.4 Taxonomy

In our work, the symbology 2525C [3] is used to describe the links between the different classification sets C_{MTI} and C_{video}. Figure 1 represents a short part of the 2525C used in this paper. The red elements underlined in italic style are the atomic elements of our taxonomy. Each element of both sets can be placed in 1. For example, the "wheeled vehicle" of the set C_{MTI} is placed at the level "Armoured $\rightarrow$ Wheeled" or the "Volkswagen Touareg" given by the video is placed at the levels "Armoured $\rightarrow$ Wheeled$\rightarrow$ Medium" and "Civilan Vehicle $\rightarrow$ Jeep $\rightarrow$ Medium".

3 Tracking with road constraints

3.1 VS IMM with a road network

The IMM is an algorithm for combining state estimates arising from multiple filter models to get a better global state estimate when the target is under maneuvers. In section 2.1, a constrained motion model i to a road segment s, noted $M^i_s(k)$, was defined. We extend the segment constraint to the different dynamic models (among a set of $r+1$ motion models) that a target can follow. The model indexed by $r=0$ is the stop model. It is evident that when the target moves from one segment to the next, the set of dynamic models changes according to the road network configuration. The steps of the IMM under road segment s constraint are the same as for the classical IMM as described in [1].

In real applications, the predicted state could also appear onto another road segment, because of a road turn for example, and we need to introduce new constrained motion models.

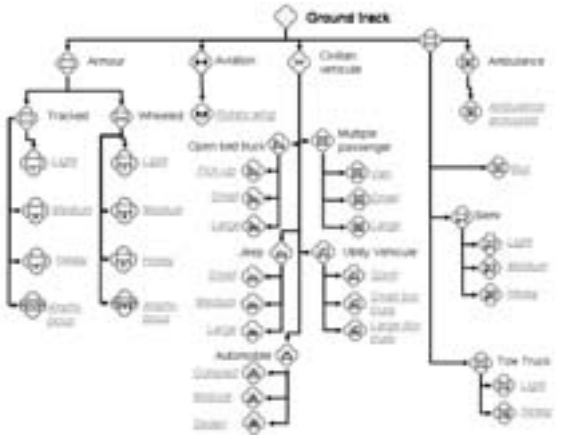


Figure 1: 2525C (light version).

In such case, we activate the most probable road segments sets depending on the local predicted statelocation of the track $T^{k,l}$[6]. We consider $r+1$ oriented graphs which depend on the road network topology. For each graph i, $i = 0, 1, \ldots, r$, each node is a constrained motion model M_s^i. The nodes are connected to each other according to the road network configuration and one has a finite set of $r+1$ motion models constrained to a road section. The selection of the most probable motion model set, to estimate the road section on which the target is moving on, is based on Wald's sequential probability ratio test (SPRT) [9].

3.2 Multiple target tracking

For the MGT problem, we use the SB-MHT (Structured Branching Multiple Hypotheses Tracking) presented in [2]. When the new measurements set $\mathbf{Z}(k)$ is received, a standard gating procedure is applied in order to validate MTI reports to track pairings. The existing tracks are updated with VS-IMMC and the extrapolated and confirmed tracks are formed. More details can be found in chapter 16 of [2]. In order to palliate the association problem, we need a probabilistic expression for the evaluation of the track formation hypotheses that includes all aspects of the data association problem. It is convenient to use the log-likelihood ratio (LLR) $L^l(k)$ or a track score of a track $T^{k,l}$ expressed at current time t_k.

4 Target type tracking

Our approach consists to use the belief on the identification attribute to revise the LLR with the *posterior* pignistic probability on the target type. We recall briefly the Target Type

Tracking (TTT) principle and explain how to improve VS-IMMC SB-MHT with target ID information. TTT is based on the sequential combination (fusion) of the predicted belief of the type of the track with the current "belief measurement" obtained from the target classifier decision. The adopted combination rule is the so-called Proportional Conflict Redistribution rule no 5 (PCR5) developed in the DSmT (Dezert-Smarandache Theory) framework since it deals efficiently with (potentially high) conflicting information. A detailed presentation with examples can be found in [4, 7].

4.1 Principle of the target type tracker

To estimate the true target type $type(k)$ at time k from the sequence of declarations $c(1)$, $c(2), \ldots c(k)$ done by the unreliable classifier up to time k. To build an estimator $\widehat{type}(k)$ of $type(k)$, we use the general principle of the Target Type Tracker (TTT) developed in [4] which consists in the following steps:

1. Initialization step (*i.e.* $k = 0$). Select the target type frame $C_{Tot} = \{\theta_1, \ldots, \theta_n\}$ and set the prior bba $m^-(.)$ as vacuous belief assignment, i.e $m^-(\theta_1 \cup \ldots \cup \theta_n) = 1$ since one has no information about the first observed target type.

2. Generation of the current bba $m_{obs}(.)$ from the current classifier declaration $c(k)$ based on attribute measurement. At this step, one takes $m_{obs}(c(k)) = P\{c(k)\} = C_{c(k)c(k)}$ and all the unassigned mass $1 - m_{obs}(c(k))$ is then committed to total ignorance $\theta_1 \cup \ldots \cup \theta_n$. $C_{c(k)c(k)}$ is the element of the known confusion matrix $\mathbf{C}$ of the classifier indexed by $c(k)c(k)$.

3. Combination of current bba $m_{obs}(.)$ with prior bba $m^-(.)$ to get the estimation of the current bba $m(.)$.

4. Estimation of True Target Type is obtained from $m(.)$ by taking the singleton of Θ, *i.e.* a Target Type, having the maximum of belief (or eventually the maximum Pignistic Probability).

5. Set $m^-(.) = m(.)$; do $k = k + 1$ and go back to step 2).

Naturally, in order to revise the LLR in our GMTI-MTT system for taking into account the estimation of belief of target ID coming from the Target Type Trackers, we transform the resulting bba $m(.) = [m^- \oplus m_{obs}](.)$ available at each time k into a probability measure.

4.2 Data attributes in the VS IMMC

To improve the target tracking process, the introduction of the target type probability is done in the likelihood calculation. For this, we consider the measurement $\mathbf{z}_j^*(k)(\forall j \in$

$\{1, \ldots, m_k\}$) described in (1) and (2). With the assumption that the kinematic and classification observations are independant, it is easy to prove that the new combined likelihood Λ_N^l associated with a track $T^{k,l}$ is the product of the kinematic likelihood.

5 Illustration

In the extended version of this paper, we will illustrate our algorithm by using a complex scenario generated with a powerful simulator developed at ONERA. The area of interest is located in a fictive country called North Atlantis. In this scenario, the goal is to detect and track several targets with 2 GMTI sensors (JSTARS, SIDM), 18 UGS and 4 UAV (SDTI), in oder to build the situation assessment and evaluate the threat in order to protect the coalition forces. On the operation theater, 250 targets evolve, they can maneuver on and out the road network. The set of target type is significant, we can have for instance civilian vehicles (as 4x4, cars, bus, truck,...) and military vehicles as well (T−62, AMX 30, Kamakov,...). llustrations and conclusion of our algorithm will be presented in the extended version of this paper.

References

[1] Y. Bar-Shalom and D. Blair, *Multitarget multisensor tracking : Applications and Advances*, Vol. III, Artech House, pp. 523, 2000.

[2] S.S. Blackman and R. Popoli, *Design and analysis of modern tracking systems*, Artech House, 1999.

[3] Defense Information Systems Agency, *Common Warfighting Symbology, MIL-STD 2525C*, Technical document, IPSC, Nov. 2008.

[4] J. Dezert, A. Tchamova, F. Smarandache and P.Konstantinova, *Target Type Tracking with PCR5 and Dempster's rules: A Comparative Analysis*, in Proc. of Fusion 2006, Firenze, Italy, July 2006.

[5] NATO, *STANAG 4607 JAS (Edition 2) - NATO ground moving target indicator GMTI) format*, NSA0749(2007)-JAS/4607, Aug. 2007.

[6] B. Pannetier, V. Nimier and M. Rombaut, *Multiple ground target tracking with a GMTI sensor*, in Proc. of MFI 2006, Sept. 2006.

[7] F. Smarandache, J. Dezert, *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3, American Research Press, 2004–2009. http://www.gallup.unm.edu/˜smarandache/DSmT.htm

[8] M. Ulmke, W. Koch, *Road-map assisted ground moving target tracking*, IEEE Trans. on AES, Vol. 42, No. 4, pp. 1264–1274, Oct. 2006.

[9] A. Wald, *Sequential Tests of Statistical Hypotheses*, Annals of Mathematical Statistics, Vol. 16, No. 2, pp. 117–186, June 1945.

Edge Detection in Color Images Based on DSmT

Jean Dezert
Zhun-ga Liu
Grégoire Mercier

Originally published as Dezert J., Liu Z., Mercier G., *Edge Detection in Color Images Based on DSmT*, in Proc. Of Fusion 2011 Conf., Chicago, July, 2011, and reprinted with permission.

***Abstract*—In this paper, we present a non-supervised methodology for edge detection in color images based on belief functions and their combination. Our algorithm is based on the fusion of local edge detectors results expressed into basic belief assignments thanks to a flexible modeling, and the proportional conflict redistribution rule developed in DSmT framework. The application of this new belief-based edge detector is tested both on original (noise-free) Lena's picture and on a modified image including artificial pixel noises to show the ability of our algorithm to work on noisy images too.**

Keywords: Edge detection, image processing, DSmT, DST, fusion, belief functions.

I. Introduction

Edge detection is one of most important tasks in image processing and its application to color images is still subject to a very strong interest [8], [10]–[12], [14] for example in teledetection, in remote sensing, target recognition, medical diagnosis, computer vision and robotics, etc. Most of basic image processing algorithms developed in the past for gray-scale images have been extended to multichannel images. Edge detection algorithms for color images have been classified into three main families [15]: 1) fusion methods, 2) multidimensional gradient methods and 3) vector methods depending on the position of where the recombination step applies [7]. In this paper, the method we propose uses a fusion method with a multidimensional gradient method. Our new unsupervised edge detector combines the results obtained by gray-scale edge detectors for individual color channels [3] to define bba's from the gradient values which are combined using Dezert-Smarandache Theory [17] (DSmT) of plausible and paradoxical reasoning for information fusion. DSmT has been proved to be a serious alternative to well-known Dempster-Shafer Theory of mathematical evidence [16] specially for dealing with highly conflicting sources of evidences. Some supervised edge detectors based on belief functions computed from gaussian pdf assumptions and Dempster-Shafer Theory can be found in [1], [21]. In this work, we show through very simple examples how edge detection can be performed based on DSmT fusion techniques with belief functions without learning (supervision). The interest for using belief functions for edge detection comes from their ability to model more adequately uncertainties with respect to the classical probabilistic modeling approach, and to deal with conflicting information due to spatial changes in the image or noises. This paper is organized as follows: In section 2 we briefly recall the basics of DSmT and the fusion rule we use. In section 3, we present in details our new edge detector based on belief functions and their fusion. Results of our new algorithm tested on the original Lena's picture and its noisy version are presented in section 4 with a comparison to the classical Canny's edge detector. Conclusions and perspectives are given in section 5.

II. Basics of DSmT

The purpose of DSmT [17] is to overcome the limitations of DST [16] mainly by proposing new underlying models for the frames of discernment in order to fit better with the nature of real problems, and proposing new efficient combination and conditioning rules. In DSmT framework, the elements θ_i, $i = 1, 2, \ldots, n$ of a given frame Θ are not necessarily exclusive, and there is no restriction on θ_i but their exhaustivity. The hyper-power set D^Θ in DSmT, the hyper-power set is defined as the set of all composite propositions built from elements of Θ with operators $\cup$ and $\cap$. For instance, if $\Theta = \{\theta_1, \theta_2\}$, then $D^\Theta = \{\emptyset, \theta_1, \theta_2, \theta_1 \cap \theta_2, \theta_1 \cup \theta_2\}$. A (generalized) basic belief assignment (bba for short) is defined as the mapping $m : D^\Theta \to [0, 1]$. The generalized belief and plausibility functions are defined in almost the same manner as in DST. More precisely, from a general frame Θ, we define a map $m(.) : D^\Theta \to [0, 1]$ associated to a given body of evidence $\mathcal{B}$ as

$$m(\emptyset) = 0 \qquad \text{and} \qquad \sum_{A \in D^\Theta} m(A) = 1 \tag{1}$$

The quantity $m(A)$ is called the *generalized* basic belief assignment/mass (or just "bba" for short) of A.

The *generalized* credibility and plausibility functions are defined in almost the same manner as within DST, i.e.

$$\text{Bel}(A) = \sum_{\substack{B \subseteq A \\ B \in D^\Theta}} m(B) \quad \text{and} \quad \text{Pl}(A) = \sum_{\substack{B \cap A \neq \emptyset \\ B \in D^\Theta}} m(B) \tag{2}$$

Two models[1] (the free model and hybrid model) in DSmT can be used to define the bba's to combine. In the free DSm model, the sources of evidence are combined without taking into account integrity constraints. When the free DSm model does not hold because the true nature of the fusion problem under consideration, we can take into account some known integrity constraints and define bba's to combine using the proper hybrid DSm model. All details of DSmT with

[1]Actually, Shafer's model, considering all elements of the frame as truly exclusive, can be viewed as a special case of hybrid model.

many examples can be easily found in [17] available freely on the web. In this paper, we will work only with Shafer's model of the frame where all elements θ_i of Θ are assumed truly exhaustive and exclusive (disjoint) and therefore D^Θ reduces the the classical power set 2^Θ and generalized belief functions reduces to classical ones as within DST framework. Aside offering the possibility to work with different underlying models (not only Shafer's model as within DST), DSmT offers also new efficient combination rules based on proportional conflict redistribution (PCR rules no 5 and no 6) for combining highly conflicting sources of evidence. In DSmT framework, the classical pignistic transformation $BetP(.)$ is replaced by the by the more effective $DSmP(.)$ transformation to estimate the subjective probabilities of hypotheses for decision-making support once the combination of bba's has been obtained. Before presenting our new edge detector, we just recall briefly what are the PCR5 fusion rule and the DSmP transformation. All details, justifications with examples on PCR5 and DSmP can be found freely from the web in [17], Vols. 2 & 3 and will not be reported here.

A. PCR5 fusion rule

The Proportional Conflict Redistribution Rule no. 5 (PCR5) is used generally to combine bba's in DSmT framework. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. Let $m_1(.)$ and $m_2(.)$ be two independent[2] bba's, then the PCR5 rule is defined as follows (see [17], Vol. 2 for full justification and examples): $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1)m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)}\right] \quad (3)$$

where all denominators in (3) are different from zero. If a denominator is zero, that fraction is discarded. Additional properties of PCR5 can be found in [5]. Extension of PCR5 for combining qualitative bba's can be found in [17], Vol. 2 & 3. All propositions/sets are in a canonical form. A variant of PCR5, called PCR6 has been proposed by Martin and Osswald in [17], Vol. 2, for combining $s > 2$ sources. The general formulas for PCR5 and PCR6 rules are given in [17], Vol. 2 also. PCR6 coincides with PCR5 when one combines two sources. The difference between PCR5 and PCR6 lies in the way the proportional conflict redistribution is done as soon as three or more sources are involved in the fusion. For example, let's consider three sources with bba's $m_1(.), m_2(.)$ and $m_3(.)$, $A \cap B = \emptyset$ for the model of the frame Θ, and $m_1(A) = 0.6$, $m_2(B) = 0.3$, $m_3(B) = 0.1$. With PCR5 the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR5} = 0.01714$ and $x_B^{PCR5} = 0.00086$ because the proportionalization requires

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B)m_3(B)}$$

$$\text{that is} \quad \frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{0.03} = \frac{0.018}{0.6 + 0.03} \approx 0.02857$$

$$\text{thus} \quad \begin{cases} x_A^{PCR5} = 0.60 \cdot 0.02857 \approx 0.01714 \\ x_B^{PCR5} = 0.03 \cdot 0.02857 \approx 0.00086 \end{cases}$$

With the PCR6 fusion rule, the partial conflicting mass $m_1(A)m_2(B)m_3(B) = 0.6 \cdot 0.3 \cdot 0.1 = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR6} = 0.0108$ and $x_B^{PCR6} = 0.0072$ because the PCR6 proportionalization is done as follows:

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_{B,2}^{PCR6}}{m_2(B)} = \frac{x_{B,3}^{PCR6}}{m_3(B)} = \frac{m_1(A)m_2(B)m_3(B)}{m_1(A) + m_2(B) + m_3(B)}$$

that is

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_{B,2}^{PCR6}}{0.3} = \frac{x_{B,3}^{PCR6}}{0.1} = \frac{0.018}{0.6 + 0.3 + 0.1} = 0.018$$

thus

$$\begin{cases} x_A^{PCR6} = 0.6 \cdot 0.018 = 0.0108 \\ x_{B,2}^{PCR6} = 0.3 \cdot 0.018 = 0.0054 \\ x_{B,3}^{PCR6} = 0.1 \cdot 0.018 = 0.0018 \end{cases}$$

and therefore with PCR6, one gets finally the following redistributions to A and B:

$$\begin{cases} x_A^{PCR6} = 0.0108 \\ x_B^{PCR6} = x_{B,2}^{PCR6} + x_{B,3}^{PCR6} = 0.0054 + 0.0018 = 0.0072 \end{cases}$$

From the implementation point of view, PCR6 is simpler to implement than PCR5. Very basic Matlab codes for PCR5 and PCR6 fusion rules can be found in [17], [18].

B. DSmP transformation

DSmP probabilistic transformation is a serious alternative to the classical pignistic transformation which allows to increase the probabilistic information content (PIC), i.e. to reduce Shannon entropy, of the approximated subjective probability measure drawn from any bba. Justification and comparisons of $DSmP(.)$ w.r.t. $BetP(.)$ and to other transformations can be found in details in [6], [17], Vol. 3, Chap. 3. $DSmP$ transformation is defined[3] by $DSmP_\epsilon(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$DSmP_\epsilon(X) = \sum_{Y \in 2^\Theta} \frac{\sum_{\substack{Z \subseteq X \cap Y \\ |Z|=1}} m(Z) + \epsilon \cdot |X \cap Y|}{\sum_{\substack{Z \subseteq Y \\ |Z|=1}} m(Z) + \epsilon \cdot |Y|} m(Y) \quad (4)$$

[2] I.e. each source provides its bba independently of the other sources.

[3] Here we work on classical power-set, but DSmP can be defined also for working with other fusion spaces, hyper-power sets or super-power sets if necessary.

where $|X \cap Y|$ and $|Y|$ denote the cardinals of the sets $X \cap Y$ and Y respectively; $\epsilon \geq 0$ is a small number which allows to increase the PIC value of the approximation of $m(.)$ into a subjective probability measure. Usually $\epsilon = 0$, but in some particular degenerate cases, when the $DSmP_{\epsilon=0}(.)$ values cannot be derived, the $DSmP_{\epsilon>0}$ values can however always be derived by choosing ϵ as a very small positive number, say $\epsilon = 1/1000$ for example in order to be as close as we want to the highest value of the PIC. The smaller ϵ, the better/bigger PIC value one gets. When $\epsilon = 1$ and when the masses of all elements Z having $|Z| = 1$ are zero, $DSmP_{\epsilon=1}(.) = BetP(.)$, where the pignistic transformation $BetP(.)$ is defined by [19]:

$$BetP\{X\} = \sum_{Y \in 2^\Theta} \frac{|Y \cap X|}{|Y|} m(Y) \tag{5}$$

with convention $|\emptyset|/|\emptyset| = 1$.

C. DS combination rule

Dempster-Shafer (DS) rule of combination is the main historical (and still widely used) rule proposed by Glenn Shafer in his milestone book [16]. Very passionate debates have emerged in the literature about the justification and the behavior of this rule from the famous Zadeh's criticism in [22]. We don't plan to reopen this endless debate and just want to recall briefly here how it is mathematically defined. Let's consider a given discrete and finite frame of discernment $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ of exclusive and exhaustive hypotheses (a.k.a satisfying Shafer's model) and two independent bba's $m_1(.)$ and $m_2(.)$ defined on 2^Θ, then DS rule of combination is defined by $m_{DS}(\emptyset) = 0$ and $\forall X \neq \emptyset$ and $X \in 2^\Theta$:

$$m_{DS}(X) = \frac{1}{1 - K_{12}} \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) \tag{6}$$

where $K_{12} \triangleq \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2)$ represents the total conflict between sources. If $K_{12} = 1$, the sources of evidence are in full conflict and DS rule cannot be applied. DS rule is commutative and associative and can be extented for the fusion of $s > 2$ sources as well. The main criticism about such such concerns its unexpected/counter-intuitive behavior as soon as the degree of conflict between sources becomes high (see [17], Vol.1, Chapter 5 and references therein for details and examples).

D. Decision-making support

Decisions are achieved by computing the expected utilities of the acts using either the subjective/pignistic $BetP\{.\}$ (usually adopted in DST framework) or $DSmP(.)$ (as suggested in DSmT framework) as the probability function needed to compute expectations. Usually, one uses the maximum of the pignistic probability as decision criterion. The maximum of $BetP\{.\}$ is often considered as a prudent betting decision criterion between the two other decision strategies (max of plausibility or max. of credibility which appears to be respectively too optimistic or too pessimistic). It is easy to show that $BetP\{.\}$ is indeed a probability function (see [19], [20]) as well as $DSmP(.)$ (see [17], Vol.2). The max of $DSmP(.)$ is considered as more efficient for practical applications since $DSmP(.)$ is more informative (it has a higher PIC value) than $BetP(.)$ transformation.

III. Edge detection based on DSmT and fusion

In this work, we use the most common RGB (Red-Green-Blue) representation of the digital color image where each layer (channel) R, G and B consists in a matrix of $n_i \times n_j$ pixels. The discrete value of each pixel in a given color channel is assumed in a given absolute interval of color intensity $[c_{min}, c_{max}]$. The principle of our new Edge detector based on DSmT is very simple and consists in the following steps:

A. Step 1: Construction of bba's

Let's consider a given channel (color layer) and denote it as L which can represent either the Red (R) color layer, the Green (G) color layer or the Blue (B) color layer, or any other channel in a more general case for multispectral images. For simplicity, we focus our work and presentation here on color images only.

Apply an edge detector algorithm for each color channel L to get for each pixel x_{ij}^L, $i = 1, 2, \ldots, n_i$, $j = 1, 2, \ldots, n_j$ an associated bba $m_{ij}^L(.)$ expressing the local belief that this pixel belongs or not to an edge. The frame of discernment Θ used to define the bba's is very simple and is defined as

$$\Theta = \{\theta_1 \triangleq \text{Pixel} \in \text{Edge}, \theta_2 \triangleq \text{Pixel} \notin \text{Edge}\} \tag{7}$$

Θ is assumed to satisfy Shafer's model (i.e. $\theta_1 \cap \theta_2 = \emptyset$). It is clear that many (binary) edge detection algorithms are available in the image processing literature but here we want a "smooth" algorithm able to provide both the belief of each pixel to belong or not to an edge and also the uncertainty one has on the classification of this pixel. In the this subsection, we present a very simple algorithm for accomplishing this task at the color channel level. Obviously the quality of the algorithm used in this first step will have a strong impact of the final result and therefore it is important to focus research efforts on the development of efficient algorithms for realizing this step as best as possible.

As in Sobel method [9], two 3×3 kernels are convolved with the original image A^L for each layer L to calculate approximations of the derivatives - one for horizontal changes, and one for vertical. We then obtain two gradient images G_x^L and G_y^L for each layer L represent the horizontal and vertical derivative approximations for each pixel x_{ij}^L. The x-coordinate is defined as increasing in the right-direction, and the y-coordinate is as increasing in the down-direction. At each pixel x_{ij}^L of the color layer L, the gradient magnitude g_{ij}^L can be estimated by the combination of the two gradient approximations as:

$$g_{ij}^L = \frac{1}{\sqrt{2}} (G_x^L(i,j)^2 + G_y^L(i,j)^2)^{1/2} \tag{8}$$

where

$$G_x^L = \frac{1}{8}\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} * A^L;$$

$$G_y^L = \frac{1}{8}\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} * A^L;$$

and where $*$ denotes the 2-dimensional convolution operation.

In Sobel's detection method, the edge detection for a pixel x_{ij} of a gray image is declared based on a hard thresholding of g_{ij} value. Such Sobel detector is sensitive to noise and it can generate false alarms. In this work, g_{ij}^L values are used only to define the mass function (bba) of each pixel in each layer over the power-set of Θ defined in (7). If the value g_{ij}^L value of a pixel is big, it implies that this pixel is more likely to belong to an edge. If g_{ij}^L value of the pixel x_{ij}^L is low then our belief that it belongs to an edge must be low too. Such very simple and intuitive modeling can be obtained directly from the sigmoid functions commonly used as activation function in neural networks, or as fuzzy membership in the fuzzy subsets theory as explained below.

Let's consider the sigmoid function defined as

$$f_{\lambda,t}(g) \triangleq \frac{1}{1+e^{-\lambda(g-t)}} \tag{9}$$

g is the gradient magnitude of the pixel under consideration. t is the abscissa of the inflection point of the sigmoid which can be selected by $t = p \cdot \max(g)$ where p is a proportion parameter and $\cdot$ is the scalar product operator. When working with noisy images, p always increases with the level of noise. λ is the slope of the tangent at the inflection point.

It can be easily verified that the bba $m_{ij}^L(.|g_{ij}^L)$ satisfying the expected behavior can be obtained by the fusion[4] of the two following simple bba's defined by:

focal element	$m_1(.)$	$m_2(.)$
θ_1	$f_{\lambda,t_e}(g)$	0
θ_2	0	$f_{-\lambda,t_n}(g)$
$\theta_1 \cup \theta_2$	$1-f_{\lambda,t_e}(g)$	$1-f_{-\lambda,t_n}(g)$

with $0 < t_n < t_e < 255, \lambda > 0$.

t_e is the lower threshold for the edge detection, and t_n is the upper threshold for the non edge detection. Thus, $[t_n, t_e]$ corresponds to our uncertainty decision zone and the g_{ij}^L values lying in this interval correspond to the unknown decision state. The bounds (thresholds) t_n and t_e can be tuned based on the average gradients values of the image, and the length $t_e - t_n$ depends on the level of the noise. If the the image is very noisy, it means the information is very uncertain, and the length of the interval $[t_n, t_e]$ can become large. Otherwise, it is small. Because of structure of these two simple bba's, the fusion obtained with PCR5, DS of even with DSm hybrid (DSmH) rules of combination provide globally similar results and therefore the choice of the fusion rule here does not really matter to build $m_{ij}^L(.|g_{ij}^L)$ as shown on the figures 1-3. PCR5, which is the most specific fusion rule (it reduces the level of belief committed to the uncertainty), is used in this work to generate $m_{ij}^L(.|g_{ij}^L)$.

[4]with DS, PCR5 or even with DSmH rule [17].

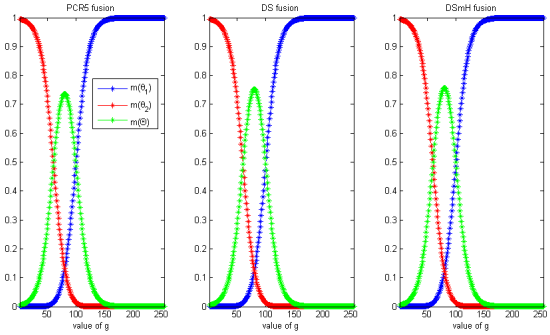


Figure 1. Computation of $m_{ij}^L(.|g_{ij}^L)$ from $m_1(.)$ and $m_2(.)$ with $[t_n, t_e] = [60, 100]$ and $\lambda = 0.09$.

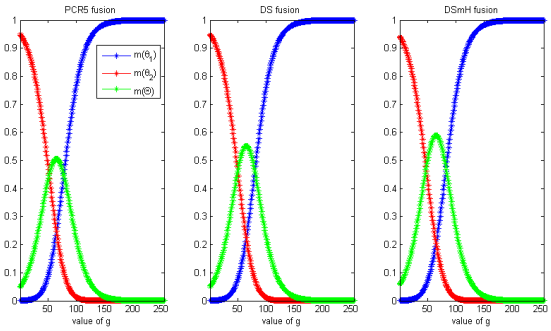


Figure 2. Computation of $m_{ij}^L(.|g_{ij}^L)$ from $m_1(.)$ and $m_2(.)$ with $[t_n, t_e] = [50, 80]$ and and $\lambda = 0.06$.

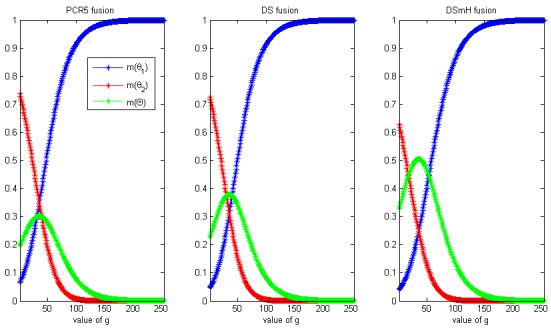


Figure 3. Computation of $m_{ij}^L(.|g_{ij}^L)$ from $m_1(.)$ and $m_2(.)$ with $[t_n, t_e] = [30, 40]$ and $\lambda = 0.04$.

In summary, $m_{ij}^L(.|g_{ij}^L)$ can be easily constructed from the choice of thresholding parameters t_e, t_n defining the uncertainty zone of the gradient values, the slope parameter λ of sigmoids, and of course from the gradient magnitude g_{ij}^L. This approach is very easy to implement and very flexible since it depends on the parameters which are totally under the control of the user.

B. Step 2: Fusion of bba's $m_{ij}^L(.)$

Many combination rules like DS rule, Dubois & Prade rule Yager's rule, and so on can be used with our approach. In this work, we just make investigations based on the two most well-known rules (DS and PCR5 rule proposed in DST and DSmT respectively). So we use either DS or PCR5 rule to combine the three bba's $m_{ij}^R(.)$, $m_{ij}^G(.)$ and $m_{ij}^B(.)$ for each pixel x_{ij} in order to get the global bba $m_{ij}(.)$ to estimate the degree of belief of the belonging of x_{ij} to an edge in the given image. Since PCR5 is not associative, we must apply the general PCR5 formula for combining the 3 sources (channels) altogether[5] as explained in details in [17], Vol.2, Chap. 1 & 2. A suboptimal approach requiring less computations would consist in applying a PCR5 sequential fusion of these bba's in such a way that the two least conflicting bba's are combined at first by PCR5 and then combine again with PCR5 the resulting bba's with the third one according to (3). The more simple PCR6 rule could also be used instead of PCR5 as well - see [17], Vol. 2.

C. Step 3: Decision-making

The output of step 2 is the set of $N_i \times N_j$ bba's $m_{ij}(.)$ associated to each pixel x_{ij} of the image in the whole color space (R,G,B). $m_{ij}(.)$ commits some degree of belief to $\theta_1 \triangleq$ Pixel $\in$ Edge, to $\theta_2 \triangleq$ Pixel $\notin$ Edge and also to the uncertainty $\theta_1 \cup \theta_2$. The binary decision-making process consists in declaring if the pixel x_{ij} under consideration belongs or not to an edge from the bba $m_{ij}(.)$, or in a more complicated manner from $m_{ij}(.)$ and the bba's of its neighbours. In this paper, we just recall the principal methods based on the use of $m_{ij}(.)$.

Based on $m_{ij}(.)$ only, how to decide θ_1 or θ_2? Many approaches have been proposed in the literature for answering this question when working with a n-D frame Θ. The pessimistic approach consists in declaring the hypothesis $\theta_i \in \Theta$ which has the maximum of credibility, whereas the optimistic approach consists in declaring the hypothesis which has the maximum of plausibility. When the cardinality of the frame Θ is greater than two, these two approaches can yield to a different final decision. In our particular application and since our frame Θ has only two elements, the final decision will be the same if we use the max of credibility or the max of plausibility criterion. Other decision-making methods suggest, as a good balance between aforementioned pessimistic and optimistic approaches, to approximate the bba at first into a subjective probability measure from a suitable probabilistic transformation, and then to choose the element of Θ which has the highest probability. In practice, one suggests to take as final decision the argument of the max of $BetP(.)$ or of the max of $DSmP(.)$. In our binary frame case however these two approaches also provide the same final decision as with the max of credibility approach. This can be easily proved from $BetP(.)$ or $DSmP(.)$ formulas. Indeed, let's consider $m(\theta 1) > m(\theta_2) > 0$ with $m(\theta_1) + m(\theta_2) + m(\theta_1 \cup \theta_2) = 1$ (which means that θ_1 is taken as final decision because it has a higher credibility than θ_2), then one gets as approximate subjective probabilities:

$$BetP(\theta_1) = m(\theta_1) + m(\theta_1 \cup \theta_2)/2 \equiv m(\theta_1) + K$$

$$BetP(\theta_2) = m(\theta_2) + m(\theta_1 \cup \theta_2)/2 \equiv m(\theta_2) + K$$

$$DSmP(\theta_1) = m(\theta_1)[1 + \frac{m(\theta_1 \cup \theta_2)}{m(\theta_1) + m(\theta_2)}] \equiv m(\theta_1)[1 + K']$$

$$DSmP(\theta_2) = m(\theta_2)[1 + \frac{m(\theta_1 \cup \theta_2)}{m(\theta_1) + m(\theta_2)}] \equiv m(\theta_2)[1 + K']$$

where K and K' are two positive constants. From these expressions, one sees that if $m(\theta 1) > m(\theta_2) > 0$, then also $BetP(\theta_1) > BetP(\theta_2)$ and $DSmP(\theta_1) > DSmP(\theta_2)$ and thus the final decision based on max of $BetP(.)$ or max of $DSmP(.)$ is finally the same. Note that when $m(\theta_1) = m(\theta_2)$, no rational decision can be drawn from $m(.)$ and only a random decision procedure or ad-hoc method can be used in such particular case.

In summary, one sees that when working with a binary frame Θ, all common decision-making strategies provide the same final decision and therefore there is no interest to use a complex decision-making procedure in that case and that's why we can adopt here the max of belief as final decision-making criterion in our simulations. Note that aside the final decision and because we have $m(\theta_1 \cup \theta_2)$, we are able (if we want) also to plot the level of uncertainty related with such decision (not presented in this paper).

IV. Simulations results

In this section we present the results of our new edge detection algorithm tested on two color images for different parameter settings.

A. Test on original Lena's picture

Lena Soderberg picture is one of the most used image for testing image processing algorithms in the literature [4] and therefore we propose to test our algorithm on this reference image. This image can be found as part of the USC SIPI Image Database in their "miscellaneous" collection available at http://sipi.usc.edu/database/index.php. The original Lena's picture scan is shown on Fig. 4-(a). The figure 5-(a)–(c) shows the edge detection on each channel (layer) based on the bba's $m_{ij}^L(.|g_{ij}^L)$ in section III-A. One sees that the edges in different channels are different, and the task of our proposed algorithm

[5] i.e. a generalization of the PCR5 formula described in section II-A.

Figure 4. Lena's picture before and after noise

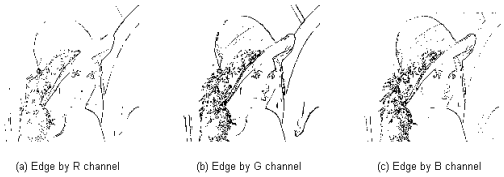


Figure 5. Edge detections in each channel.

Figure 6. Canny's edge detector on Lena's gray image.

Figure 7. Sobel's edge detector on Lena's gray image.

Figure 8. DS-based edge detector on Lena's color image.

Figure 9. PCR5 edge detector on Lena's color image.

is to combine efficiently the underlying bba's $m_{ij}^L(.|g_{ij}^L)$ generating the subfigures 5-(a)–(c).

Sobel [9] and Canny [2] edge detectors are commonly used in image processing community and that's why we make comparison of our new edge detector w.r.t. Canny's and Sobel's approaches. Canny and Sobel edge detectors are applied directly to the gray image converted from the original Lena color image Fig. 4-(a). The figures 6–9 show the results of the different edge detectors on Lena's picture. In our simulations, we took $\lambda = 0.06$, and t_g defined as $t = p \cdot \max(g)$ in each layer, was taken with $p_n = 0.17$ and $p_e = 0.19$, corresponding to gradient thresholds $[t_n^R, t_e^R] = [15, 17]$, $[t_n^G, t_e^G] = [13, 14]$ and $[t_n^B, t_e^B] = [11, 13]$. The max of credibility, plausibility, $DSmP$ or $BetP$ for decision-making to generate final result provide the same decision as explained in the section III-C which is normal in this binary frame case.

One sees that finally on the clean (noise-free) Lena's picture, our edge detector provides close performances to Sobel's detector applied on Lena's grey image. Canny's detector seems to provide a better ability to detect some edges in Lena's picture than our method, but it also generates much more false

alarms too. It is worth noting that the results provided by DS-based or PCR5-based edge detectors show a coarse location of the edges. So it is quite difficult to drawn a clear and fair conclusion between these edge detectors since it highly depends on what we want, i.e. the reduction of false alarms or the reduction of miss-detections.

B. Test on Lena's picture with noise

In this simulation, we show how our edge detector works on a noisy image. Sampling of independent Gaussian noise $\mathcal{N}(0, \sigma^2)$ is added to each pixels of each layer of the original Lena's picture as seen on Fig. 4-(b). In the presented simulation, $\sigma^2 = 1100$ which correspond approximatively to the value of the variance of the blue channel and half the variance of the others. Local edge detection for each layer based on $m_{ij}^L(.|g_{ij}^L)$ is shown on Fig. 10-(a)–(c), where the red points represent the ignorant pixel which commits the most belief to the ignorance $\theta_1 \cup \theta_2$. As shown in Fig 10, the edge detection in each channel is very noisy. Our method allows to commit automatically highest belief value to uncertainty for most of pixels associated to an edge which actually correspond to noises.[6] The edge detection based on fusion result are interesting as shown by Fig.11 and Fig. 12 because it shows the ability of our edge detector to suppress the noise effects. For comparison, we give on Fig. 13 and Fig.14, the performance of Canny and Sobel edge detectors applied classically on the noisy gray-level Lena's picture. In this simulation, we took $\lambda = 0.06$, and t using $p_n = 0.22 \cdot \max(g)$ and $p_e = 0.39 \cdot \max(g)$ in each layer with $[t_n^R, t_e^R] = [36, 20]$, $[t_n^G, t_e^G] = [35, 19]$ and $[t_n^B, t_e^B] = [31, 18]$. The decision-making is still based on max of credibility.

The visual comparison and analysis of results shown of figures 11–12 clearly indicates that our edge detector based on the fusion of belief constructed on each layer works much better than the edge detection applied separately on each layer. There is no ignorant pixel corresponding to red color according to the fusion results, since the fusion process of DS or PCR5 rule effectively decrease the uncertainty. Our results show also clearly that Canny and Sobel edge detectors applied to noisy gray-level Lena's picture are very sensitive to the noise perturbations. Our proposed method (based on DS rule or on PCR5 rule) is more robust to the noise perturbations and provides better results than Sobel or Canny edge detector for such noisy image. For this tested image, it appears that the results using DS and PCR5 rules are very close, because there is not too much conflict actually between bba's of layers and one know that in such case PCR5 rule behavior is close to DS rule behavior. DS rule is usually good enough in the low conflict case, whereas PCR5 rule is preferred for the combination of high conflicting sources of evidence. So the preference of PCR5 with respect to DS rule for edge detection must be guided by the level of conflict which appears in the layers of the color image that we need to process.

[6] So we are also able at layer level to filter these pixels (false alarms) before applying the fusion. This has not yet be done in this work.

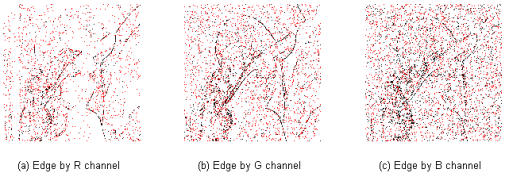


Figure 10. Edge detections in each channel on noisy image.

Figure 11. DS edge detector on noisy Lena's color image.

V. Conclusions and Perspectives

A new unsupervised edge detector for color image based on belief functions has been proposed in this work. The basic belief assignment (bba) associated with the edge of a pixel in each channel of the image is defined according to its gradient magnitude, and one can easily model the uncertainty about our belief it belong or not to an edge. PCR5 and DS rules have been applied in this work to combine these bba's to get the global bba for final decision-making. Other rules of combination of bba's could also have been used instead but they are known to be less efficient than PCR5 or DS rules in high and low conflict cases respectively. The

Figure 12. PCR5 edge detector on noisy Lena's color image.

Figure 13. Sobel's edge detector on noisy Lena's gray image.

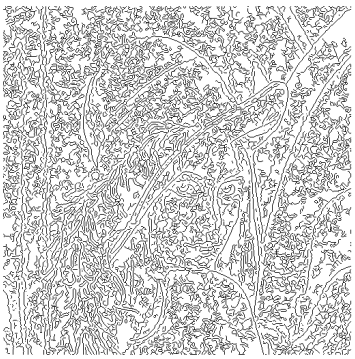


Figure 14. Canny's edge detector on noisy Lena's gray image.

fusion process is able to reduce noise perturbations because the noises are assumed to be independent between channels. The final decision making on the edge can be made either on the maximum of credibility, plausibility, $DSmP$ or $BetP$ values as well. The first simulation done on original Lena's picture shows that our edge detector works as well as the classical Sobel's edge detector and it provides less false alarms than with Canny's detector, but seems to generate more miss-detections. In our second simulation based on noisy Lena image, the results show that our new edge detector is more robust to the noise perturbations than Sobel or Canny classical edge detectors. As possible improvement of this algorithm and for further research, we would like to include some morphological or connexity constraints at a higher level of processing and develop automatic technique for threshold selection. The application of this new approach of edge detection to satellite multispectral images is under investigations.

REFERENCES

[1] S. Ben Chaabane, F. Fnaiech, M. Sayadi, E. Brassart, *Relevance of the Dempster-Shafer evidence theory for image segmentation*, 3rd International conference on Signals, Circuits and Systems, Nov. 6-8th, Medenine, Tunisia, 2009.

[2] J.F. Canny, *A Computational Approach to Edge Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6):679-698, 1986.

[3] C.J. Delcroix, M. A. Abidi, *Fusion of edge maps in color images*, in Proc. SPIE Int. Soc. Opt. Eng. 1001, pp. 454–554, 1988.

[4] N. Devillard, *Image Processing: the Lena story*, 2006. http://ndevilla.free.fr/lena/, see also http://www.lenna.org.

[5] J. Dezert, F. Smarandache, *Non Bayesian conditioning and deconditioning*, Int. Workshop on Belief Functions, Brest, France, April 2010.

[6] J. Dezert, F. Smarandache, *A New Probabilistic Transformation of Belief Mass Assignment (DSmP)*, in Proc. of Fusion 2008 Int. Conf., Cologne, Germany, June 30-July 3, 2008.

[7] A.N. Evans, *Nonlinear Edge Detection in color images*, Chap. 12, pp. 329-356, in [13] .

[8] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1991.

[9] E.P. Lyvers and OR. Mitchell, *Precision Edge Contrast and Orientation Estimation*, IEEE Transactions on Pattern Analysis and Machine Intelligence,10(6):927-937,1988.

[10] A. Koschan, M. Abidi, *Detection and classification of edges in color images*, IEEE Signal Processing Magazine, vol. 22, no. 1, pp. 64–73, 2005.

[11] A. Koschan, M. Abidi, *Digital Color Image Processing*, John Wiley Press, NJ, USA, 376 pages, May 2008.

[12] D. Marr, E. Hildreth, *Theory of Edge Detection*, in Proceedings of the Royal Society London 207, pp. 187–217, 1980.

[13] S. Marshall,G. L. Sicuranza (Editors), *Advances in nonlinear signal and image processing*, EURASIP Book series on Signal Processing and Communications, Hindawi Publishing Corporation, 2006.

[14] K. N. Plataniotis, A. N. Venetsanopoulos, *Color Image Processing and Applications*, Springer, New York, NY, USA, 2000.

[15] M. A. Ruzon, C. Tomasi, *Edge, junction, and corner detection using color distributions*, IEEE Transactions on Pattern Analaysis and Machine Intelligence, vol. 23, no. 11, pp. 1281–1295, 2001.

[16] G. Shafer, *A Mathematical Theory of Evidence*, Princeton Univ. Press, 1976.

[17] F. Smarandache, J. Dezert (Editors), *Advances and Applications of DSmT for Information Fusion*, American Research Press, Rehoboth, Vol.1-3, 2004-2009.

[18] F. Smarandache, J. Dezert, J.-M. Tacnet, *Fusion of sources of evidence with different importances and reliabilities*, in Proceedings of Fusion 2010 conference, Edinburgh, UK, July 2010.

[19] P. Smets, *Constructing the pignistic probability function in a context of uncertainty*, Uncertainty in Artificial Intelligence, Vol. 5, pp. 29–39, 1990.

[20] P. Smets, R. Kennes, *The transferable belief model*, Artif. Intel., 66(2), pp. 191-234, 1994.

[21] P. Vannoorenberghe, O. Colot, D. De Brucq, *Color Image Segmentation Using Dempster-Shafer's Theory*, ICIP 99. Proc. of 1999 International Conference on Image Processing, Kobe, Japan, Vol. 4, pp. 300–304, 1999.

[22] L. Zadeh, *On the validity of Dempster's rule of combination*, Memo M79/24, Univ. of California, Berkeley, USA, 1979.

DSmT Applied to Seismic and Acoustic Sensor Fusion

Erik P. Blasch
Jean Dezert
Pierre Valin

Originally published as Blasch E., Dezert J., Valin P., *DSmT Applied to Seismic and Acoustic Sensor Fusion,* Proc. IEEE Nat. Aerospace Electronics Conf (NAECON), 2011, and reprinted with permission.

Abstract – *In this paper, we explore the use of DSMT for seismic and acoustic sensor fusion. The seismic/acoustic data is noisy which leads to classification errors and conflicts in declarations. DSmT affords the redistribution of masses when there is a conflict. The goal of this paper is to present an application and comparison on DSMT with other classifier methods to include the support vector machine(SVM) and Dempster-Shafer methods. The work is based on two key references (1) Marco Duarte with the initial SVM classifier application of the seismic and acoustic sensor data and (2) Arnaud Martin in Vol. 3 with the Proportional Conflict Redistribution Rule 5/6 (PCR5/PCR6) developments. By using the developments of Duarte and Martin, we were able to explore the various aspects of DSMT in an unattended ground sensor scenario. Using the receiver operator curve (ROC), we compare the methods for individual classification as well as a measure of overall classification using the area under the curve(AUC). Conclusions of the work show that the DSMT affords a lower false alarm rate because the conflict information is redistributed over the set masses and is comparable to other classifier results when using a maximum decision forced choice.*

Keywords: *Information Fusion, DSMT, PCR5, PCR6, Area Under the Curve(AUC), SVM.*

1 Introduction

The goal of this paper is to present an application and comparison on DSMT with other classifier methods. The work is based on two key references (1) Marco Duarte with the initial classifier application of the seismic and acoustic sensor data [**1**] and (2) Arnaud Martin in Vol. 3 for the implementation of the DSMT methods. [**2**] By using the developments of Duarte and Martin, we were able to explore the various aspects of DSMT in an unattended ground sensor scenario. In the exploration of information fusion metrics for classification, there is a need to develop metrics of effectiveness that support the user's utility needs [**3**] and can vary over the sensor types, environmental conditions, targets of interest, situational context, and users [**4**].

DSmT is an extension to the Dempster-Shafer method of evidential reasoning which has been detailed in numerous papers and texts: *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3 [**5**].

In 2002, Dezert [**6**] introduced the methods for the reasoning and in 2003, presented the hyper power-set notation for DSmT [**7**]. Recent applications include the DSmT Proportional Conflict Redistribution rule 5 (PCR5) applied to target tracking [**8**].

The key contributions of DSmT are the redistributions of masses such that no refinement of the frame Θ is possible unless a series of constraints are known. For example, Shafer's model [**9**] is the DSm hybrid model in DSmT. Since Shafer's model, authors have continued to refine the method to more precisely address the combination of conflicting beliefs [**10**, **11**, **12**] and generalization of the combination rules [**13**, **14**]. An adaptive combination rule [**15**] and rules for quantitative and qualitative combinations [**16**] have been proposed. Recent examples for sensor applications include electronic support measures, [**17**, **18**] and physiological monitoring sensors [**19**]. One application of DSmT that has not been fully explored is in seismic, magnetic, and acoustic classification fusion of moving targets. Kadambe conducted an information theory approach [**20**] and used DSmT as integrity constraints [**21**], but did not take advantage of the conflict redistribution.

Detecting moving vehicles in an urban area [**22**] is an example where DSmT conflicting mass redistribution could be helpful [**8**]. Detecting traffic can be completed by fixed ground cameras or on dynamic unattended ground vehicles (UGVs). If the sensors are on UGVS, path planning is needed to route the UGVs to observe the traffic [**23**, **24**] and cooperation among UGVs is necessary[**25**]. The DARPA Grand Challenge featured sensors on mobile UGVs observing the environment [**26**]. Mobile sensing can be used to orient [**27**] or conduct simultaneous location and mapping (SLAM) [**28**].

Deployed ground sensors can observe the vehicles; however they are subject to the quality of the sensor measurements as a well as obscurations. One interesting question is how to deploy the fixed sensors that optimize the performance of a system. Efforts in distributed wireless networks (WSNs) have resulted in many issues in distributed processing, communications, and data fusion [**29**]. In a dynamic scenario, resource coordination [**30**] is needed for both context assessment, but also the ability to be aware of impending situational threats [**31**, **32**]. For distributed sensing systems, to combine sensors, data, and user analysis requires pragmatic approaches to metrics [**33**, **34**, **35**, **36**]. For example, Zahedi [**37**] develops a

QOI architecture for comparison of centralized versus distributed sensor network deployment planning.

Information fusion has been interested in the problems of databases for target trafficability[**38**], sensor management [**39**], and processing algorithms [**40**] from which to assess objects in the environment. Various techniques have incorporated grouping object movements [**41**], road information [**42**, **43**], and updating the object states based on environmental constraints [**44**]. Detecting, classifying, identifying and tracking objects [**45**] has been important for a variety of sensors, including 2D visual, radar [**46**], and hyperspectral [**47**] data; however newer methods are of interest for ground sensors with 1D signals.

Seismic data provides passive sensing of ground vibrations which can be used for motion tracking. Passive magnetic sensing can detect hidden objects that might indicate intent. Finally, acoustic data can be used for signature detection from vehicle engines. [**48**] The DARPA SENSIT program investigated deploying a distributed set of wireless sensors along a road to classify vehicles as shown in **Figure 1**.

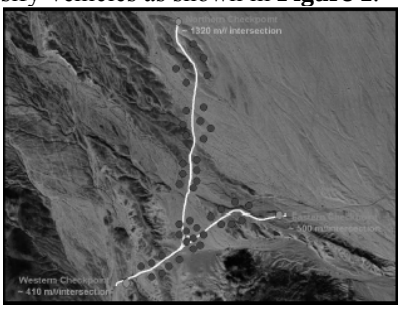


Figure 1. SENSIT Data from [M. F. Duarte and Y. H. Hu, "Vehicle Classification in Distributed Sensor Networks," 2004 [**49**]

The sensors include acoustic and seismic signals. Given the deployed set of sensors, feature vectors were used to classify signals based on the data from the seismic and acoustic signals. [**49**] Various approaches include combining the data with decision fusion [**50**], value fusion [**51**], and simultaneous track and identification methods [**52**, **53**]. Information theoretical approaches including the KL method were applied to the data for sensor management [**54**] as shown in **Figure 2**. Processing sensor data for target classification using acoustic [**55**, **56**] and seismic [**57**] results have been explored in support of information theoretical sensor placement [**58**].

Much work has been completed using imaging sensors and radar sensors for observing and tracking targets. Video sensors are limited in power and subject to day/night conditions. Likewise, radar line-of site precludes them from observing in the same plane. Together, both imaging and radar sensors do not have the advantage of UGSs which can power on and off, can work for a long time on battery power, and can be deployed to remote areas.

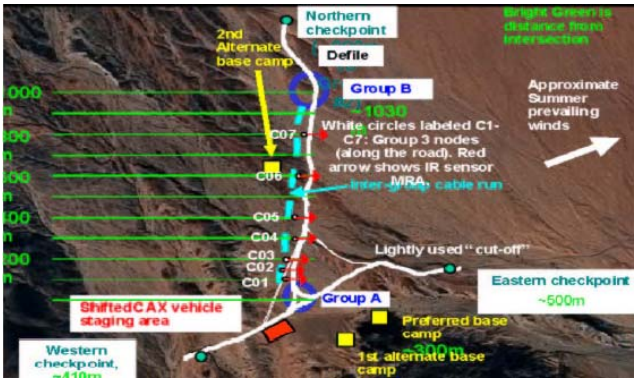


Figure 2. Deployed Sensors. From S. Kadambe and C. Daniell, "Theoretic Based Performance of Distributed Sensor Networks", *AFRL-IF-RS-TR-2003*, 231, October 2003. [**54**]

Track management situational awareness tools receive input from sensor feeds (examples include electro-optical, radar, electronic support measures (ESMs), and sonar) and display this information to a user. User inputs include: creation of new objects, such as tracks, contacts and targets. Methods to reduce data-to-decisions include: fusing multiple tracks into a single track, incorporating alerting mechanisms, or visualizing track data common operational picture (COP). Sensor and track data can grow rapidly as the user desires to keep historical data.

Our goal is to utilize the DSmT method for the fusion of information from seismic and acoustic data in which each sensor/classifier is in direct conflict with the other sensor. We address (1) intelligent use of the data based on value for classification, (2) DSMT sensor data fusion for detection, classification, and positional location, and (3) metrics to support the sensor and data management as supporting a user control.

2 Location / Detection

We desire to track and identify the targets based on the sensor reports. In this study, we concentrate on the classification of targets which can be used with the kinematic/position information for target identification.

2.1 Sensor Information Management

The goal is to utilize the UGSs sensors which may be acoustic, magnetic, seismic, and PIRoelectric (passive infrared for motion detection. With a variety of sensors, information fusion can (a) utilize the most appropriate sensor at the correct time, (b) combine information from both sensors on a single platform, (c) combine results from multiple platforms, and (d) cue other sensors in a hand-off fashion to effectively monitor the area. Sensor exploitation requires an analysis of feature generation, extraction, and selection or (construction, transformation, selection, and evaluation). To provide track and ID results, we develop method or target classification.

2.2 Sensor Classification

Sensor exploitation includes detection, recognition, classification, identification and characterization of some object. Individual classifiers can be deployed at each level to robustly determine the object information. Popular methods include voting, neural networks, fuzzy logic, neuro-dynamic programming, support vector machines, Bayesian and Dempster-Shafer methods. One way to ensure the accurate assessment is to look at a combination of classifiers. Combination of classifiers [**59**] could include different sensors with classifiers, different methods over a single or multiple sensors, and various hierarchies of coordinating the classifiers such as Bayes nets and distributed processing.

Issues in classifier combination methods need to be compared as related to decisions, feature sets, and user involvement. Selecting the optimal feature set is based on the situation and environmental context of which the sensors are deployed. An important question for sensor and data management is measures of effectiveness. For instance, what is the quantification of fusion/decision gain using a set of classification methods and placement methods? There is a need for a robust combination rule that includes the location and detection of the sensors subject to the target and environmental constraints. Typically, a mobile sensor needs to optimize its route and can be subject to interactive effects of pursuers and evaders with other targets [**60**] as well as active jamming of the signal [**61**].

Detecting targets from seismic and acoustic data in a distributed net centric fashion requires pragmatic approaches to sensor and data management. [**62**] To robustly track and ID a target requires both the structured data from the kinematic movements as well as the unstructured data for the feature analysis. [**63**]

3 DSMT

Here we use PCR6 and PCR5 and the DSMP selections which are discussed below. We replace Smets' rule [**10**] by the more effective Proportional Conflict Redistribution rule no. 5 (PCR5) or eventually the more simple PCR rule no. 6 (PCR6) and replace the pignistic transformation by the more effective DSmP transformation to estimate target classification probabilities. All details, justifications with examples on PCR5 and PCR6 fusion rules and DSmP transformation can be found freely from the web in the DSmT compiled texts [**5**], Vols. 2 & 3..

3.1 PCR5 and PCR6 fusion rules

In DSmT (Dezert-Smarandache Theory) framework, the Proportional Conflict Redistribution Rule no. 5 (PCR5) is used generally to combine the basic belief assignment (bba)'s. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. Let $m_1(.)$ and $m_2(.)$ be two independent bba's, then the PCR5 rule is defined as follows (see [**5**], Vol. 2 for full justification and examples): $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)} \right]$$

where all denominators in the equation above are different from zero. If a denominator is zero, that fraction is discarded. Additional properties of PCR5 can be found in [**64**]. Extension of PCR5 for combining qualitative bba's can be found in [**5**], Vol. 2 & 3. All propositions/sets are in a canonical form. A variant of PCR5, called *PCR6* has been proposed by Martin and Osswald in [**5**], Vol. 2, for combining $s > 2$ sources. PCR6 coincides with PCR5 when one combines two sources. The difference between PCR5 and PCR6 lies in the way the proportional conflict redistribution is done as soon as three or more sources are involved in the fusion. For example, let's consider three sources with bba's $m_1(.)$, $m_2(.)$, and $m_3(.)$, $A \cap B = \emptyset$ for the model of the frame Θ, and $m_1(A) = 0.6$, $m_2(B) = 0.3$, and $m_3(B) = 0.1$. With **PCR5** the partial conflicting mass $m_1(A)\ m_2(B)\ m_3(B) = (0.6)(0.3)(0.1) = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR5} = 0.01714$ and $x_B^{PCR5} = 0.00086$ because the proportionalization is [**8**]:

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)\ m_3(B)} = \frac{m_1(A)\ m_2(B)\ m_3(B)}{m_1(A) + m_2(B)\ m_3(B)}$$

that is $$\frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{(0.3)(0.1)} = \frac{0.018}{0.6 + 0.03} \approx 0.02857$$

thus $$x_A^{PCR5} = 0.60\ (0.02857) \approx 0.01714$$
$$x_B^{PCR5} = 0.03\ (0.02857) \approx 0.00086$$

With the **PCR6 fusion rule**, the partial conflicting mass $m_1(A)\ m_2(B)\ m_3(B) = (0.6)(0.3)(0.1) = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR6} = 0.0108$ and $x_B^{PCR6} = 0.0072$ because the PCR6 proportionalization is done as follows:

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_{B;2}^{PCR6}}{m_2(B)} = \frac{x_{B;3}^{PCR6}}{m_3(B)} = \frac{m_1(A)\ m_2(B)\ m_3(B)}{m_1(A) + m_2(B) + m_3(B)}$$

that is

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_{B;2}^{PCR6}}{0.3} = \frac{x_{B;3}^{PCR6}}{0.1} = \frac{0.018}{0.6 + 0.3 + 0.1} \approx 0.018$$

thus

$$x_A^{PCR6} = (0.6)\,(0.018) = 0.0108$$
$$x_{B,2}^{PCR6} = (0.3)\,(0.018) = 0.0054$$
$$x_{B,3}^{PCR6} = (0.1)\,(0.018) = 0.0018$$

and therefore with PCR6, one gets finally the following redistributions to A and B:

$$x_A^{PCR6} = (0.6)\,(0.018) = 0.0108$$

$$x_B^{PCR6} = x_{B,2}^{PCR6} + x_{B,3}^{PCR6} = 0.0054 + 0.0018 = 0.0072$$

From the implementation point of view, PCR6 is simpler to implement than PCR5. For convenience, Matlab codes of PCR5 and PCR6 fusion rules can be found in [5].

3.2 The DSmP Transformation

The DSmP probabilistic transformation is an alternative to the classical pignistic transformation which allows us to increase the probabilistic information content (PIC), i.e. to minimize the Shannon entropy, of the approximated subjective probability measure drawn from any bba. Justification and comparisons of DSmP(.) with respect to BetP(.) and to other transformations can be found in details in [**65**, **5** Vol. 3, Chap. 3].

BetP: The pignistic transformation probability, denoted BetP, offers a compromise between maximum of credibility *Bel* and maximum of plausibility *Pl* for decision support. The *BetP* transformation is defined by BetP(∅) = 0 and $\forall X \in G^\Theta \setminus \{\emptyset\}$ by

$$BetP(X) = \sum_{Y \in G^\Theta} \frac{\mathcal{C}_\mathcal{M}(X \cap Y)}{\mathcal{C}_\mathcal{M}(Y)} m(Y)$$

where G^Θ corresponds to the hyper-power set including all the integrity constraints of the model (if any). $G^\Theta = 2^\Theta$ if one adopts Shafer's model for Θ and $G^\Theta = D^\Theta$ (Dedekind's lattice) if one adopts the free DSm model for Θ [5]. CM(*Y*) denotes the DSm cardinal of the set *Y*, which is the number of parts of *Y* in the Venn diagram of the model M of the frame Θ under consideration [5, Book 1, Chap. 7]. The BetP reduces to the Transferable Belief Model (TBM) when G^Θ reduces to classical power set 2^Θ when one adopts Shafer's model.

DSmP transformation is defined by $DSmP_\epsilon(\emptyset) = 0$ and $\forall X \in G^\Theta \setminus \{\emptyset\}$ by

$$DSmP_\epsilon(X) = \sum_{Y \in G^\Theta} \frac{\sum_{\substack{Z \subseteq X \cap Y \\ \mathcal{C}(Z)=1}} m(Z) + \epsilon \cdot \mathcal{C}(X \cap Y)}{\sum_{\substack{Z \subseteq Y \\ \mathcal{C}(Z)=1}} m(Z) + \epsilon \cdot \mathcal{C}(Y)} m(Y)$$

where C(*X* ∩ *Y*) and C(*Y*) denote the cardinals of the sets *X* ∩ *Y* and *Y* respectively; ε ≥ 0 is a small number which allows to reach a highest PIC value of the approximation of *m*(.) into a subjective probability measure. Usually ε = 0, but in some particular degenerate cases, when the $DSmP_{\varepsilon=0}(.)$ values cannot be derived, the $DSmP_{\varepsilon>0}$ values can however always be derived by choosing ε as a very small positive number, say ε = 1/1000 for example in order to be as close as we want to the highest value of the PIC. The smaller ε, the better/bigger PIC value one gets. When ε = 1 and when the masses of all elements *Z* having C(*Z*) = 1 are zero, $DSmP_{\varepsilon=1}(.) = BetP(.)$.

4 Example/Simulation

We use the SENSIT data which was described above and was provides an unstructured data analysis. To perform the data management we use data mining [66] techniques such as a *support vector machine* (SVM) [67, 68] to process the unstructured data. Through analysis, we can determine the optimum use of the data given environmental conditions (i.e. obscurations) and sensor's capabilities to detect a moving target.

Figure 3 shows the methodology of comparison. A key comparison is made between combining all the acoustic and seismic data together for testing and training via the SVM versus using the outputs from the acoustic and seismic data separately from which conflicts in classification are detected and sent to DS and DSmT processing.

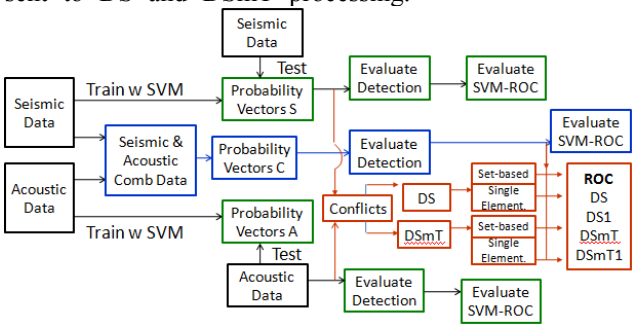


Figure 3. Experimentation Flow.

4.1 Data Processing

We compare two cases of (1) processing the data separately and (2) jointly processing the acoustic and seismic results **Figure 4** shows the case of the acoustic results.

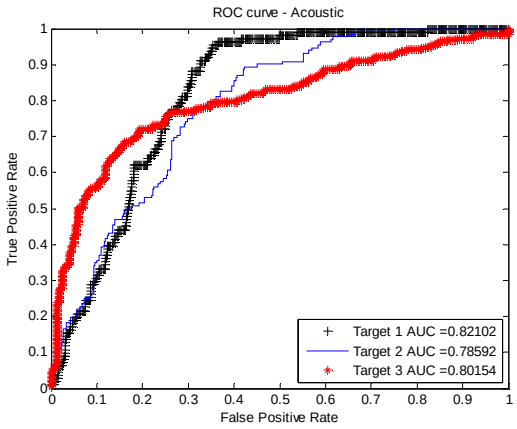


Figure 4. Acoustic Results.

Figure 5 demonstrates the results for the seismic results. Note that for the data set, the seismic results have a lower probability of false alarms for target 3 and target 2; however, target 2 exhibits more confusion.

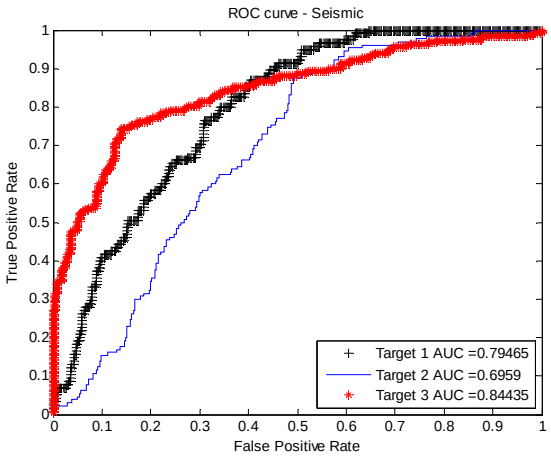


Figure 5. Seismic Results.

Next we explore the case of the joint seismic and acoustic data management and utilize SVM for classification, shown in **Figure 6**. Note the false alarm reduction which is desired by users.

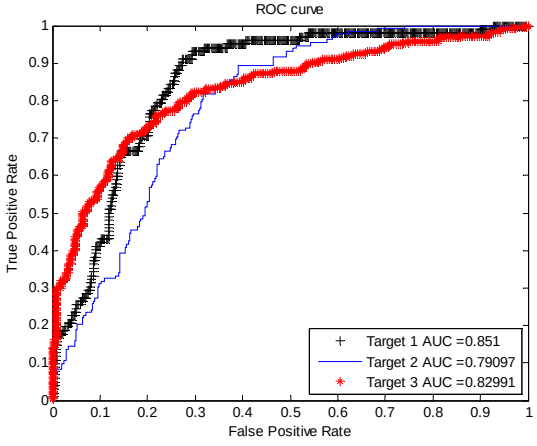


Figure 6. Combined Results

In general, the joint analysis supports better decision making as confidence was PD was improved for a constant false alarm rate, accuracy was improved as to the target location from joint spatial measurements, and timeliness in decision making as fewer measurements were needed to confirm the target ID (i.e. decision made with two modalities required fewer measurements than that of a single modality).

4.2 Application of DS

Below, we show the results of the application of DS methods. Given a training and prediction results in a combined probability, we have for target1, target2, and target3 a vector of $\mathbf{P} = [P_1 P_2 P_3]$. Based on the prediction results from the SVM, there are many conflicts of the sensor decision based on the maximum probability. When a conflict occurs, it would be better suited to acknowledge the conflict and then redistribute the probabilities based on a set notation. In this case, the focal elements are $\Phi = [\theta_1, \ldots, \theta_7] = [$'1', '1⋂2', '2', '2⋂3', '3', '1⋂3', '1⋂2⋂3'$]$. Using the analysis by Martin, we conduct an analysis over the set criterion. **Figure 7** shows that a significant reduction false alarms; however, the overall classification as measured by the area under the curve (AUC) is less than that of the SVM by itself. Thus, there is a trade off when using DS for reducing the PD for low FA versus the overall classification analysis.

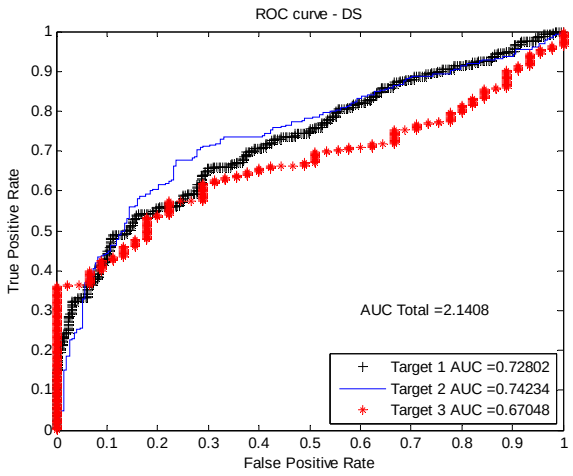


Figure 7. DS allowing for set declarations.

To explore a comparison of approaches, we utilized the bba and forced the evidential reasoned to choose a single target. From this analysis, the AUC improves in comparison to the SVM approaches which are a forced choice analysis. **Figure 8** plots the DS (for one target designation).

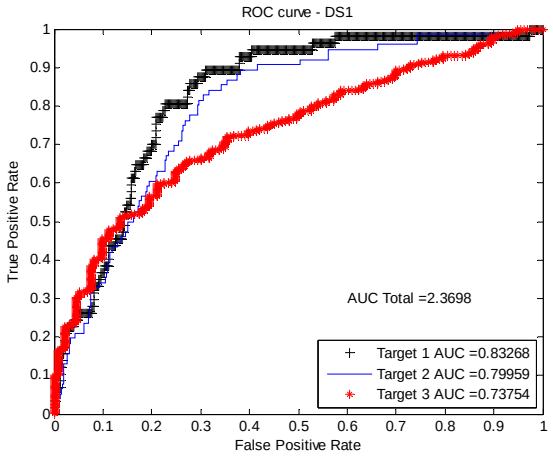


Figure 8. DMST Single Target Detection.

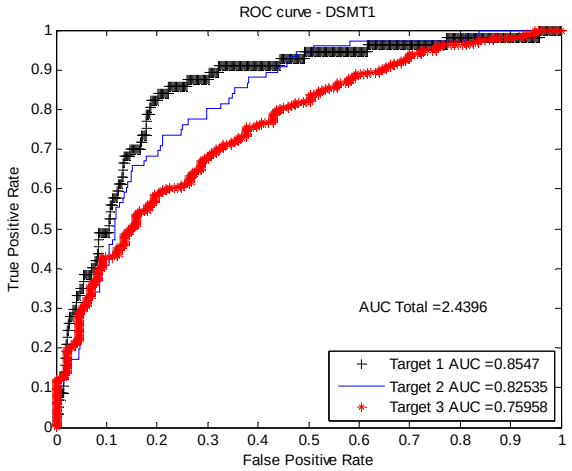


Figure 10. DMST Single Target Detection.

4.3 Application of DSMT

DMST, as described above, improves on the methods of conflict redistribution. In this case, there were slight alterations in the bba comparisons; however with the heuristic logic, changes resulted in the classification that was comparable to the complete SVM fusion analysis.

Figure 9 presents the DSMT results for set declaration and **Figure 10** shows the case of a forced target choice from the DSMT. From these plots, we can see that the set-based approach improves the detection for low false alarm rates; however for high false alarm rates, the detection probability is increased over all false alarm rates. Using the maximum of the target bba provides an analysis threshold that renders the DSmT comparable to a SVM (which is allowed to train over all the data available).

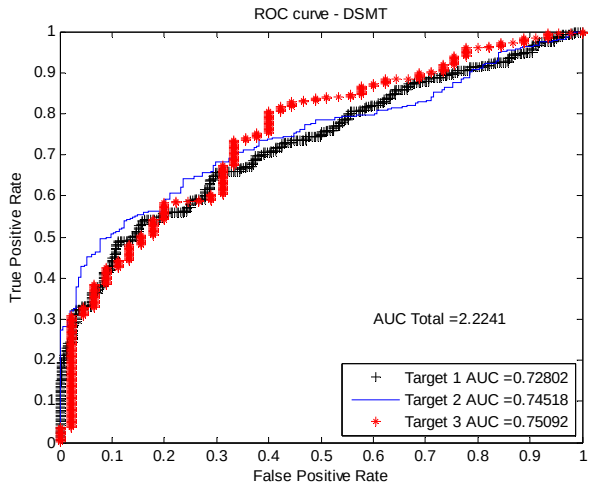


Figure 9. DSMT allowing for set declarations.

In the table below, we look at the entire analysis using the area under the ROC (AUC) as a key metric in the analysis. Additionally, there are cases in which the maximum AUC and minimum AUC are improved but the overall analysis (Total AUC) varies. We see from the comparison that the DS and DSMT methods can improve single target detection; however the SVM alone (run over all the data) does perform slightly higher in the information fusion case.

Table 1: AUC Comparisons of SVM, DS, and DSMT

Method	**Min AUC**	**Max AUC**	**Total AUC**
A-SVM	0.786	0.821	2.401
S-SVM	0.696	0.844	2.335
C-SVM	**0.791**	**0.851**	**2.472**
DS	0.671	0.742	2.141
DS1	0.738	0.833	2.371
DSMT	0.728	0.751	2.224
DSMT1	**0.760**	**0.855**	**2.440**

A – Acoustic, S-Seismic, C-Combination

5 Conclusions

We have explored DS and DSMT methods for seismic and acoustic information fusion. The goal of the paper was a new application of the existing techniques presented by Martin and Durate for further demonstration of the various modifications to the DS methods. Using the initial results, the use of DSMT can be tailored to the seismic and acoustic sensors which demonstrate high conflicts in decision outputs as they measure different target phenomenologies. We utilized a Bayesian basic belief assignment (bba) with only singleton as focal elements which from the P vectors of the target probabilities. Future work will use non-Bayesian approaches to get the bbas.

Information theoretic measures [**69**] and tracking analysis [**70**] can support the sensor and data management as well as determine the Quality of Information and Quality of Service needs. Use of the *Area Under the*

Curve (AUC) provides decision support for situational awareness for command and control from which we can extend to higher dimensions [**71**]. Various other sources of soft data (human reports) can be combined with the hard (physics-based sensing) [**72**] to update the sensor management, placement, and reporting of the situation based on the context and the needs of users such as measures of effectiveness for mission support.

Acknowledgement

Advance in research is accelerated with the helpfulness of previous researchers. We would like to thank **Marco. F. Durate** for the correspondence in the data and preliminary code. The key to the paper is the documentation and code from **Arnaud Martin** for which he his code is freely available on the web at http://bfas.iutlan.univ-rennes1.fr/news.html.

6 References

[1] M. F. Duarte and Y. H. Hu, "Vehicle Classification in distributed sensor networks," *J. Parallel Distributed Computation*, 64. No7, 2004.

[2] A. Martin and C. Osswald, "Generalized Proportion conflict redistribution rule applied to sonar imagery and radar targets classification," in *Adv. and Appl. of DSmT for Information Fusion*, Vol. 2, (Eds.) J. Dezert and F. Smarandache, 2006.

[3] E. P. Blasch "Assembling an Information-fused Human-Computer Cognitive Decision Making Tool," *IEEE Aerospace and Electronic Systems Magazine,* June 2000.

[4] E. Blasch, P. Valin, and E. Bosse, "Measures of Effectiveness for High-Level Fusion", *Fusion10,* 2010.

[5] J. Dezert and F. Smarandache, *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3, American Research Press, 2004–2009. http://fs.gallup.unm.edu/DSmT.htm

[6] J. Dezert**,** "Foundations for a new theory of plausible and paradoxical reasoning", *Information & Security, An Int'l J.*, ed. by Prof. Tzv. Semerdjiev, CLPP, Bulg. Acad. of Sciences, Vol. 9, 2002.

[7] J. Dezert and F. Smarandache**,** "On the generation of hyper-powersets for the DSmT," *Proc. Fusion03*, July 2003.

[8] J. Dezert and B. Pannetier, "A PCR-BIMM filter for maneuvering target tracking," *Fusion10*, 2010.

[9] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton Univ. Press, 1976.

[10] P. Smets, "Analyzing the Combination of Conflicting Belief Functions, 2005.

[11] M. Daniel, "The MinC Combination of Belief Functions, derivation and formulas," *Tech Report No. 964*, Acad. Sci of the Czech Republic, June 2006.

[12] A. Jøsang and M. Daniel, "Strategies for Combining Conflict Dogmatic Beliefs," *Fusion06*, 2006.

[13] M. Daniel, "Generalization of the Classic Combination Rules to DSm Hyper-Power Sets," *Information & Security, An Int'l J.*, Vol. 20, 2006.

[14] F. Smaradache and J. Dezert, "Information fusion based on new proportional conflict redistribution rules," in *Proc. 8th Int. Conf. Inf. Fusion*, 2005.

[15] M. C. Florea, J. Dezert, P. Valin, F. Smarandache, and A-L. Jousselme, "Adaptive combination rule and proportional conflict redistribution rule for information fusion," *Cogis '06 Conf.*, Paris, March 2006.

[16] A. Martin, C. Osswald, J. Dezert, and Fl. Smarandache, "General Combination Rules for Qualitative and Quantitative Beliefs," *J. Adv. Information Fusion*, Vol. 3, No. 2, Dec. 2008.

[17] P. Valin, P. Djiknavorian, and E. Bosse, "A Pragmatic Approach for the use of Dempster-Shafer Theory in Fusing realistic Sensor Data," *J. Adv. Information Fusion*, Vol. 5, No. 1, June 2010.

[18] P. Djiknavorian, D. Grenier, and P. Valin, "Approximation in DSm theory for fusing ESM reports," *Int. Workshop on Belief functions* 2010, Brest, France, April 2010.

[19] Z. H. Lee, J. S. Choir, and R. Elmasri, "A Static Evidential Network for Context Reasoning in Home-Based Care," IEEE Trans. Sys., Man, and Cyber-Part A; Sys & Humans, Vol. 40, No. 6, Nov, 2010.

[20] S. Kadambe and C. Daniell, "Sensor/Data Fusion Based on Value of Information," *Fusion03*, 2003.

[21] S. Kadambe, "Power and Resource Aware Distributed Smart Fusion," Ch. 18 in *Advances and applications of DSmT for information fusion (Collected works) Vol. 1*, 2004.

[22] H. Ling, L. Bai, E. Blasch, and X. Mei, "Robust Infrared Vehicle Tracking across Target Pose Change using L1 Regularization," *Fusion10,* 2010.

[23] E. Blasch, S. Kondor, M. Gordon, and R. Hsu, "Georgia Tech Aerial Robotics team competition entry," *J. Aerial Unmanned Vehicle Systems*, May 1994

[24] E. P. Blasch, "Flexible Vision-Based Navigation System for Unmanned Aerial Vehicles," *Proc. SPIE*. 1994.

[25] D. Shen, G. Chen, J. Cruz, and E. Blasch, "A game Theoretic Data Fusion Aided Path Planning Approach for Cooperative UAV Control," *IEEE Aerospace Conf.*, Big Sky, MT, 1 March, 2008.

[26] G Seetharaman, A. Lakhotia, E. Blasch," Unmanned Vehicles Come of Age: The DARPA Grand Challenge," *IEEE Computer Society Magazine*, Dec 2006.

[27] K. M. Lee, K.M., Z. Zhi, R. Blenis, and E. P. Blasch, "Real-time vision-based tracking control of an unmanned vehicle," *IEEE J. of Mechatronics - Intelligent Motion Control*, 1995

[28] H. Durrant-Whyte & T. Bailey, "Simultaneous Localization and Mapping (SLAM): Part I The Essential Algorithms". *Robotics and Automation Magazine* **13**: 99–110, 2006.

[29] P. K. Varshney, *Distributed Detection and Data* Fusion, Springer-Verlag, NY, 1997.

[30] E. Blasch, I. Kadar, K. Hintz, J. Biermann, C. Chong, and S. Das, "Resource Management Coordination with Level 2/3 Fusion", *IEEE AES Magazine ,* Mar. 2008.

[31] E. Blasch, I. Kadar, J. Salerno, M. M. Kokar, S. Das, G. M. Powell, D. D. Corkill, and E. H. Ruspini, "Issues and challenges of knowledge representation and reasoning methods in situation assessment (Level 2 Fusion)", *J. of Adv. in Information Fusion,* Dec. 2006.

[32] G. Chen, D. Shen, C. Kwan, J. Cruz, M. Kruger, and E. Blasch, "Game Theoretic Approach to Threat Prediction and Situation Awareness," *Advances in Journal for Information Fusion,* July 2008.

[33] C. Liu, D. Grenier, A-L. Jousselme, and E. Bosse, "Reducing Algorithm Complexity for Computing an

Aggregate Uncertainty Measure," *IEEE Trans. on Sys Man and Cybernetics - A*, Vol. 37, No. 5, Sept 2007.

[34] E. Blasch, M. Pribilski, *et. al.*, "Fusion Metrics for Dynamic Situation Analysis," *Proc SPIE Vol. 5429*, 2004.

[35] L. Bai and E. Blasch, "Two-Way Handshaking Circular Sequential k-out-of-n Congestion System," *IEEE Journal of Reliability*, Mar 2008.

[36] E. Blasch, "Derivation of a Reliability Metric for Fused Data Decision Making," *IEEE NAECON Conf.*, July 2008.

[37] S. Zahedi and C. Bisdikian, "A Framework for QOI-inspired analysis for sensor network deployment planning," *Proc Wireless Internet CON*, 2007.

[38] R. Anthony, *Principles of Data Fusion Automation*, Artech House, 1995.

[39] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.

[40] D. L. Hall and S. A. McMullen, *Mathematical Techniques in Multisensor Data Fusion*, Artech House, 2004.

[41] E. Blasch and T. Connare, "Improving Track maintenance Through Group Tracking," *Proc on Estimation, Tracking, and Fusion; Tribute to Y. Bar Shalom*, 2001.

[42] C. Yang , E. Blasch, & M Bakich, "Nonlinear Constrained Tracking of Targets on Roads", *Fusion05*, 2005.

[43] C. Yang and E. Blasch, "Fusion of Tracks with Road Constraints," *Advances in Journal for Information Fusion*, Sept. 2008.

[44] C. Yang and E. Blasch, "Kalman Filtering with Nonlinear State Constraints", *IEEE Transactions AES*, Vol. 45, No. 1, Jan. 2009.

[45] E. P. Blasch, *Derivation of a Belief Filter for Simultaneous HRR Tracking and ID*, Ph.D. Thesis, Wright State University, 1999.

[46] C. Yang, and E. Blasch, W. Garber, and R. Mitchell, "A Track-ID Maneuver Indicator from Target's Range-Doppler Image, *J. Adv. in Information Fusion*, July, 2007.

[47] T. Wang, Z. Zhu, and E. Blasch, "Bio-Inspired Adaptive Hyperspectral Imaging for real-Time Target Tracking " *IEEE Sensors Journal*, 2009.

[48] A. Bornstein, T. Damarla, J. Lavery, F. Morelli, and E. Schmeisser, "Remote Detection of Covert Tactical Adversarial Intent of Individuals in Asymmetric Operations," *ARL Tech Report ARL-SR-197* April 2010.

[49] M. F. Duarte and Y. H. Hu, "Vehicle Classification in distributed sensor networks," *J. Parallel Distributed Computation*, 64. No7, 2004.

[50] M. Duarte and Y-H Hu, "Distance Based decision Fusion in a Distributed Wireless Sensor Network" *Telecommunication Systems*, Vol. 26 No. 2-4, 2004.

[51] T. Clouqueur. P. Ramanathan, K. K. Saluja, H-C Wang, "Value-Fusion versus Decision-Fusion for Fault tolerance in Collaborative Target Detection in Sensor Networks," *Fusion01*, 2001.

[52] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, Classification and Tracking of Targets in Distributed Sensor Networks," *IEEE Sig Processing Magazine*, 2002.

[53] R. R. Brooks, P. Ramanathan, and A. M. Sayeed, "Distributed Target Classification and Tracking in Sensor Networks," *Proc of IEEE*, 2003.

[54] S. Kadambe and C. Daniell, "Theoretic Based Performance of Distributed Sensor Networks", *AFRL-IF-RS-TR-2003*, 231, October 2003.

[55] H. Xing, F. Li, H. Xiao, Y. Wang, and Y. Liu, "Ground Target Detection, Classification, and Sensor Fusion in Distributed Fiber Seismic Sensor Networks," *Proc SPIE*, vol. 6830, 2007.

[56] G. M. Jacyna, C. T. Christoi, B. George, and B. F. Necioglu, "Netted Sensor-based Vehicle Acoustic Classification at Tier 1 nodes," *Proc SPIE*, 2005.

[57] F. Valenta and H. Hermansky, "Combination of Acoustic Classifiers based on Dempster-Shafer Theory of Evidence," *IEEE ICASSP Conf*, 2007.

[58] E. P. Blasch, P. Maupin, and A-L Jousselme, "Sensor-based Allocation for path Planning and Area Coverage using UGSs", *IEEE NAECON 2010*, 2010.

[59] L. Kuncheva, *Combing Pattern Classifiers: Methods and Algorithms*, John Wiley and Sons, 2004.

[60] M. Wei, G. Chen, J. Cruz, L Haynes, M. Kruger, and E. Blasch, "A decentralized Approach to Pursuer-Evader Games with Multiple Superior Evaders in Noisy Environments," *Advances in Journal for Information Fusion*, June 2007.

[61] M. Wei, G. Chen, J Cruz, L. Haynes, M. Chang, and E. Blasch, "Confrontation of Jamming and Estimation in Pursuit-Evasion Games with Multiple Superior Evaders," AIAA *Journal of Aerospace Computing*, Aug 2006.

[62] I. Kadar, E. Blasch, and C. Yang, "Network and Service Management Effects on Distributed Net-Centric Fusion Data Quality," July, *Fusion08*, 2008.

[63] H. Chen, G. Chen, E. Blasch, and T. Schuck, "Robust Track Association and Fusion with Extended Feature Matching", invited Chapter in *Optimization & Cooperative Ctrl. Strategies*, M.J. Hirsch et al. (Eds.):, LNCIS 381, pp. 319–354, Springer-Verlag Berlin Heidelberg 2009.

[64] J. Dezert, F. Smarandache, "Non Bayesian conditioning and deconditioning," *Int. Workshop Belief Functions, Brest*, France, April 2010.

[65] J. Dezert, F. Smarandache, "A new probabilistic transformation of belief mass assignment," *Proc. Fusion08*, 2008.

[66] I. Witten and E. Frank, Data Mining: *Practical machine learning tools and techniques*, Morgan Kaufmann, San Francisco, 2005.

[67] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

[68] R.-E. Fan, P.-H. Chen, and C.-J. Lin. "Working set selection using second order information for training SVM", *Journal of Machine Learning Research* 6, 1889-1918, 2005.

[69] H. Chen, G. Chen, E. Blasch, and P. Douville, "Information Theoretic Measures for Performance Evaluation and Comparison ," *ISIF Conf. Fusion09*, 2009.

[70] O. Straka, M Flidr, J. Dunik, M, Simandl, and E. P. Blasch, "Nonlinear Estimation Framework in Target Tracking," *Fusion10*, 2010.

[71] S. Alsing, E. P. Blasch, and R. Bauer, "3D ROC Surface concepts for Evaluation of Target Recognition algorithms faced with the Unknown target detection problem," *Proc. SPIE, Vol. 3718*, Orlando, FL, 13-17 April 1999.

[72] E. P. Blasch, É. Dorion, P. Valin, E. Bossé, and J. Roy, "Ontology Alignment in Geographical Hard-Soft Information Fusion Systems," *Fusion10*, 2010.

The Effective Use of the DSmT for Multi-Class Classification

Nassim Abbas
Youcef Chibani
Arnaud Martin
Florentin Smarandache

A version of N. Abbas, Y. Chibani, and H. Nemmour, *Handwritten Digit Recognition Based On a DSmT-SVM Parallel Combination*, in Proc. 13th International Conference on Frontiers in Handwriting Recognition (ICFHR), Bari, Italy, pp. 241-246, September 18-20, 2012, and reprinted with permission.

Abstract

The extension of the Dezert-Smarandache theory (DSmT) for the multi-class framework has a feasible computational complexity for various applications when the number of classes is limited or reduced typically two classes. In contrast, when the number of classes is large, the DSmT generates a high computational complexity. This paper proposes to investigate the effective use of the DSmT for multi-class classification in conjunction with the Support Vector Machines using the One-Against-All (OAA) implementation, which allows offering two advantages: firstly, it allows modeling the partial ignorance by including the complementary classes in the set of focal elements during the combination process and, secondly, it allows reducing drastically the number of focal elements using a supervised model by introducing exclusive constraints when classes are naturally and mutually exclusive. To illustrate the effective use of the DSmT for multi-class classification, two SVM-OAA implementations are combined according three steps: transformation of the SVM classifier outputs into posterior probabilities using a sigmoid technique of Platt, estimation of masses directly through the proposed model and combination of masses through the Proportional Conflict Redistribution (PCR6). To prove the effective use of the proposed framework, a case study is conducted on the handwritten digit recognition. Experimental results show that it is possible to reduce efficiently both the number of focal elements and the classification error rate.

Keywords: *Handwriting digit recognition; Support Vector Machines; Dezert-Smarandache theory; Belief assignments; Conflict management*

1. Introduction

Nowadays a large number of classifiers and methods of generating features is developed in various application areas of pattern recognition [1,2]. Nevertheless, it failed to underline the incontestable superiority of a method over another in both steps of generating features and classification. Rather than trying to optimize a single classifier by choosing the best features for a given problem, researchers found more interesting to combine the recognition methods [2,3]. Indeed, the combination of classifiers allows exploiting the redundant and complementary nature of the responses issued from different classifiers.

Researchers have proposed various approaches for combining classifiers increasingly numerous and varied, which led the development of several schemes in order to treat data in different ways [2,3]. Generally, three approaches for combining classifiers can be considered: parallel approach, sequential approach and hybrid approach [2]. Furthermore, these ones can be performed at a class level, at a rank level, or at a measure level [4-7].

In many applications, various constraints do not allow an efficient joint use of classifiers and feature generation methods leading to an inaccurate performance. Thus, an appropriate operating method using mathematical approaches is needed,

which takes into account two notions: uncertainty and imprecision of the responses of classifiers. In general, the most theoretical advances which have been devoted to the theory of probabilities are able to represent the uncertain knowledge but are unable to model easily the information which is imprecise, incomplete, or not totally reliable. Moreover, they often lead to confuse both concepts of uncertainty and imprecision with the probability measure. Therefore, new original theories dealing with uncertainty and imprecise information have been introduced, such as the fuzzy set theory [8], evidence theory [9,10], possibility theory [11] and, more recently, the theory of plausible and paradoxical reasoning [12-14].

The evidence theory initiated by Dempster and Shafer termed as Dempster-Shafer theory (DST) [9,10] is generally recognized as a convenient and flexible alternative to the bayesian theory of subjective probability [15]. The DST is a powerful theoretical tool which has been applied in many kinds of applications [16] for the representation of incomplete knowledge, belief updating and for the combination of evidence [17,18] through the Dempster-Shafer's combination rule. Indeed, it offers a simple and direct representation of ignorance and has a low computational complexity [19] for most practical applications.

Nevertheless, this theory presents some weaknesses and limitations mainly when the combined evidence sources become very conflicting. Furthermore, the Shafer's model itself does not allow necessary holding in some fusion problems involving the existence of the paradoxical information. To overcome these limitations, a recent theory of plausible and paradoxical reasoning, known as Dezert-Smarandache theory (DSmT) in the literature, was elaborate by Jean Dezert and Florentin Smarandache for dealing with imprecise, uncertain and paradoxical sources of information. Thus, the main objective of the DSmT was to introduce combination rules that would allow to correctly combining evidences issued from different information sources, even in presence of conflicts between sources or in presence of constraints corresponding to an appropriate model (free or hybrid DSm models [12]). The DSmT has proved its efficiency in many current pattern recognition application areas such as remote sensing [20-23], identification and tracking [24-29], biometrics [30-33], computer vision [34-36], robotics [37-42] and more recently handwritten recognition applications [7,43,44] as well as many others [12-14].

The use of the DSmT for multi-class classification has a feasible computational complexity for various applications when the number of classes is limited or reduced typically two classes [43]. In contrast, when the number of classes is large, the DSmT generates a high computational complexity closely related to the number of elements to be processed. Indeed, an analytical expression defined by Tombak *et al.* [45] shows that the number of elements to be processed follows the sequence of Dedekind's numbers [46,47]: 1,2,5,19,167,7580,7828353,... For instance, if the number of classes belonging to discernment space is 8, then the number of elements to be deal in DSmT framework is $\approx 5.6\times 10^{22}$. Hence, it is not easy to consider the set of all subsets of the original classes (but under the union and the intersection operators) since it becomes untractable for more than 6 elements in the discernment space [48]. Thus, Dezert and Smarandache [49] proposed a first work for ordering all elements generated using the free DSm model for matrix calculus such as made in DST framework [50,51]. However, this proposition has limitations since in practical applications it is more appropriate to only manipulate the focal elements [7, 52-54].

Hence, few works have already been focused on the computational complexity of the combination algorithms formulated in DSmT framework. Djiknavorian and Grenier [53] showed that there's a way to avoid the high level of complexity of DSm hybrid (DSmH) combination algorithm by designing a such code that can perform a complete DSmH combination in very

short period of time. However, even if they have obtained an optimal process of evaluating DSmH algorithm, first some parts of their code are really not optimized and second it has been developed only for a dynamic fusion. Martin [55] further proposed a practical codification of the focal elements which gives only one integer number to each part of the Venn diagram representing the discernment space. Contrary to the Smarandache's codification [48] used in [56] and the proposed codes in [53], author thinks that the constraints given by the application must be integrated directly in the codification of the focal elements for getting a reduced discernment space. Therefore, this codification can drastically reduce the number of possible focal elements and so the complexity of the DST as well as the DSmT frameworks. A disadvantage of this codification is that the complexity increases drastically with the number of combined sources especially when dealing with a problem in the multi-class framework. To address this issue, Li *et al.* [57] proposed a criterion called evidence supporting measure of similarity (ESMS), which consists in selecting, among all sources available, only a subset of sources of evidence in order to reduce the complexity of the combination process. However, this criterion has been justified for only a two-class problem. Nowadays, the complexity of reducing both the number of combined sources and the size of the discernment space are research challenges that still need to be addressed.

In many pattern recognition applications, the classes belonging to the discernment space are naturally and then mutually exclusive such as in biometrics [30-33] and handwritten recognition applications [7,43,44]. Hence, several classification methods have been proposed as template matching techniques [58-60], minimum distance classifiers [61,62], support vector machine (SVM) [63], hidden Markov Models (HMMs) [63-65], neural networks [66,67]. In various pattern recognition applications, the SVMs have proved their performance from the mid-1990s comparatively to other classifiers [2]. The SVM is based on an optimization approach in order to separate two classes by an hyperplane. In the context of multi-class classification, this optimization approach is possible [68] but requiring a very costly duration. Hence, two preferable methods of multi-class implementation of SVMs have been proposed for combining several binary SVMs, , which are One Against All (OAA) and One Against One (OAO), respectively [69-71]. The former is the most commonly used implementation in the context of multi-class classification using binary SVMs, which constructs n SVMs to solve a n-class problem [72]. Each SVM is designed to separate a simple class θ_i from all the others, i.e., from the corresponding complementary class $\overline{\theta_i} = \bigcup_{\substack{0\le j\le n-1 \\ j\neq i}} \theta_j$. In contrast, the OAO implementation is designed to separate two simple classes θ_i and θ_j ($i \neq j$), which requires $n\times(n-1)/2$ SVMs. Hence, various decision functions can be used such as the Decision Directed Acyclic Graph (DDAG) [73] since it has the advantage to eliminate all possible unclassifiable data.

Generally, the combination of binary classifiers is performed through very simple approaches such as voting rule or a maximization of decision function coming from the classifiers. In this context, many combination operators can be used, especially in the DST framework [74]. Still in the same vein, some works have been tried out the combination of binary classifier originally from SVM in the DST framework [75,76]. For instance, the pairwise approach has been revisited by Quest *et al.* [76-79] in the framework of the DST of belief functions for solving a multi-class problem. In [80], the combination method based on DST has been used by Hu *et al.* for combining multiple multi-class probability SVM classifiers in order to deal with distributed multi-source multi-class problem [80]. Martin and Quidu proposed an original approach based on DST [81] for combining binary SVM classifiers using OAO or OAA strategies, which provides a decision support

helping experts for seabed characterization from sonar images. Burger *et al.* [82] proposed to apply a belief-based method for SVM fusion to hand shape recognition. Optimizing the fusion of the sub-classifications and dealing with undetermined cases due to uncertainty and doubt have been investigated by other works [83], through a simple method, which combines the fusion methods of belief theories with SVMs. Recently, one regression based approach [84] has been proposed to predict membership or belief functions, which are able to model correctly uncertainty and imprecision of data.

In this work, we propose to investigate the effective use of the DSmT for multi-class classification in conjunction with the SVM-OAA implementation, which allows offering two advantages: firstly, it allows modeling the partial ignorance by including the complementary classes in the set of focal elements, and then in the combination process, contrary to the OAO implementation which takes into account only the singletons, and secondly, it allows reducing drastically the number of focal elements from $Dedekind(n)$ to $2\times n$. The reduction is performed through a supervised model using exclusive constraints. Combining the outputs of SVMs within DSmT framework requires that the outputs of SVMs must be transformed into membership degree. Hence, several methods of estimating of mass functions are proposed in both DST and DSmT frameworks, these ones can be directly explicit through special functions or indirectly explicit through transfer models [9,85-88]. In our case, we propose a direct estimation method based on a sigmoid transformation of Platt [89]. This allows us to satisfy the OAA implementation constraint.

The paper is organized as follows. Section 2 reviews the Proportional Conflict Redistribution (PCR6) rule based on DSmT. Section 3 describes the combination methodology for multi-class classification using the SVM-OAA implementation. Experiments conducted on the dataset of the isolated handwritten digits are presented in section 4. The last section gives a summary of the proposed combination framework and looks to the future research direction.

2. Review of PCR6 combination rule

In pattern recognition, the multi-class classification problem is generally formulated as a n-class problem where classes are associated to patterns classes, namely $\theta_0,\theta_1,\ldots,$ and θ_n. Hence, the parallel combination of two classifiers, namely information sources S_1 and S_2, respectively, is performed through the PCR6 combination rule based on the DSmT. For n-class problem, a reference domain also called the discernment space should be defined for performing the combination, which is composed of a finite set of exhaustive and mutually exclusive hypotheses.

In the context of the probabilistic theory, the discernment space, namely Θ, is composed of n elements as: $\Theta=\{\theta_0,\theta_1,\ldots,\theta_n\}$, and a mapping function $m\in[0,1]$ is associated for each class, which defines the corresponding mass verifying $m(\varnothing)=0$ and $\sum_{i=0}^{n} m(\theta_i)=1$. In Bayesian framework, combining two sources of information by means of the weighted mean and consensus based rules seems effective for non-conflicting responses [90-93]. In the opposite case, an alternative approach has been developed in DSmT framework to deal with (highly) conflicting imprecise and uncertain sources of information [14]. Example of such approaches is PCR6 rule.

The main concept of the DSmT is to distribute unitary mass of certainty over all the composite propositions built from elements of Θ with $\cup$ (Union) and $\cap$ (Intersection) operators instead of making this distribution over the elementary hypothesis only. Therefore, the *hyper-powerset* D^Θ is defined as:

1. $\emptyset, \theta_0, \theta_1, \ldots, \theta_n \in D^\Theta$.
2. If $A, B \in D^\Theta$, then $A \cap B \in D^\Theta$ and $A \cup B \in D^\Theta$.
3. No other elements belong to D^Θ, except those obtained by using rules 1 or 2.

The DSmT uses generalized basic belief mass, also known as the generalized basic belief assignment (gbba) computed on hyper-powerset of Θ and defined by a map $m(.): D^\Theta \to [0,1]$ associated to a given source of evidence which can support paradoxical information, as follows: $m(\emptyset) = 0$ and $\sum_{A \in D^\Theta} m(A) = 1$. The combined masses m_{PCR6} obtained from $m_1(.)$ and $m_2(.)$ by means of the PCR6 rule [13,14] are defined as:

$$m_{PCR6}(A_i) = \begin{cases} 0 & \text{if } A_i \in \Phi, \\ m_\wedge(A_i) + \sum_{k=1}^{2} m_k^2(A_i) L_k & \text{otherwise.} \end{cases} \tag{1}$$

Where

$$L_k = \sum_{\substack{Y_{\sigma_k(1)} \cap A_i \in \Phi \\ Y_{\sigma_k(1)} \in D^\Theta}} \frac{m_{\sigma_k(1)} Y_{\sigma_k(1)}}{m_k(A_i) + m_{\sigma_k(1)} Y_{\sigma_k(1)}} \tag{2}$$

$\Phi = \{\Phi_M, \emptyset\}$ is the set of all relatively and absolutely empty elements, Φ_M is the set of all elements of D^Θ which have been forced to be empty in the hybrid model M defined by the exhaustive and exclusive constraints, $\emptyset$ is the empty set, the denominator $m_k(A_i) + m_{\sigma_k(1)} Y_{\sigma_k(1)}$ is different to zero, and where $\sigma_k(1)$ counts from 1 to 2 avoiding k, i.e.:

$$\sigma_k(1) = \begin{cases} 2 & \text{if } k = 1, \\ 1 & \text{if } k = 2. \end{cases} \tag{3}$$

Thus, the term $m_\wedge(A_i)$ represents a conjunctive consensus, also called DSm Classic (DSmC) combination rule [13,14], which is defined as:

$$m_\wedge(A_i) = \begin{cases} 0 & \text{if } A_i = \emptyset, \\ \sum_{(X, Y \in D^\Theta, X \cap Y = A_i)} m_1(X) m_2(Y) & \text{otherwise.} \end{cases} \tag{4}$$

3. Methodology

The proposed combination methodology shown in Fig. 1 is composed of two individual systems using SVMs classifiers. Each one is trained using its own source of information providing two kinds of complementary features, which are combined through the PCR6 rule. In the following, we give a description of each module composed our system.

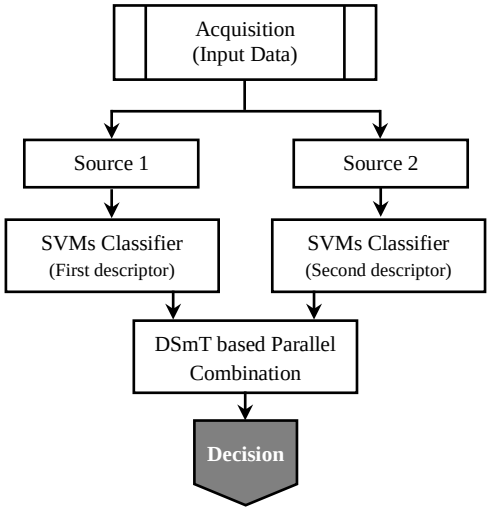


Fig 1. Structure of the combination scheme using SVM and DSmT

3.1. Classification based on SVM

The classification based on SVMs has been used widely in many pattern recognition applications as the handwritten digit recognition [2]. The SVM is a learning method introduced by Vapnik *et al*. [94], which tries to find an optimal hyperplane for separating two classes. Its concept is based on the maximization of the distance of two points belonging each one to a class. Therefore, the misclassification error of data both in the training set and test set is minimized.

Basically, SVMs have been defined for separating linearly two classes. When data are non linearly separable, a kernel function K is used. Thus, all mathematical functions, which satisfy Mercer's conditions, are eligible to be a SVM-kernel [94]. Examples of such kernels are sigmoid kernel, polynomial kernel, and Radial Basis Function (RBF) kernel. Then, the decision function $f : \mathrm{R}^{p} \rightarrow \{-1,+1\}$, is expressed in terms of kernel expansion as:

$$f(x) = \sum_{k=1}^{Sv} \alpha_k \, y_k \, K(x, x_k) + b \tag{5}$$

where α_k are Lagrange multipliers, Sv is the number of support vectors x_k which are training data, such that $0 \le \alpha_k \le C$, C is a user-defined parameter that controls the tradeoff between the machine complexity and the number of nonseparable points [73], the bias b is a scalar computed by using any support vector.

Finally, for a two-class problem, test data are classified according to:

$$x \in \begin{cases} class\,(+1) & \text{if } f(x) \succ 0 \\ class\,(-1) & \text{otherwise} \end{cases} \tag{6}$$

The extension of the SVM for multi-class classification is performed according the One Against-All (OAA) [95]. Let a set of N training samples which are separable in n classes $\{\theta_0, \theta_1, \ldots, \theta_{n-1}\}$, such that $\left\{\left(x_k^i, y_k^i\right) \in \mathrm{R}^p \times \{\pm 1\}; k = 1,..,N; i = 1,..n\right\}$. The principle consists to separate a class from other classes. Consequently, n SVMs are required for solving n class problem.

3.2. Classification Based On DSmT

The proposed classification based on DSmT is presented in Fig. 2, which is conducted into three steps: i) estimation of masses, ii) combination of masses through the PCR6 combination rule and iii) decision rule.

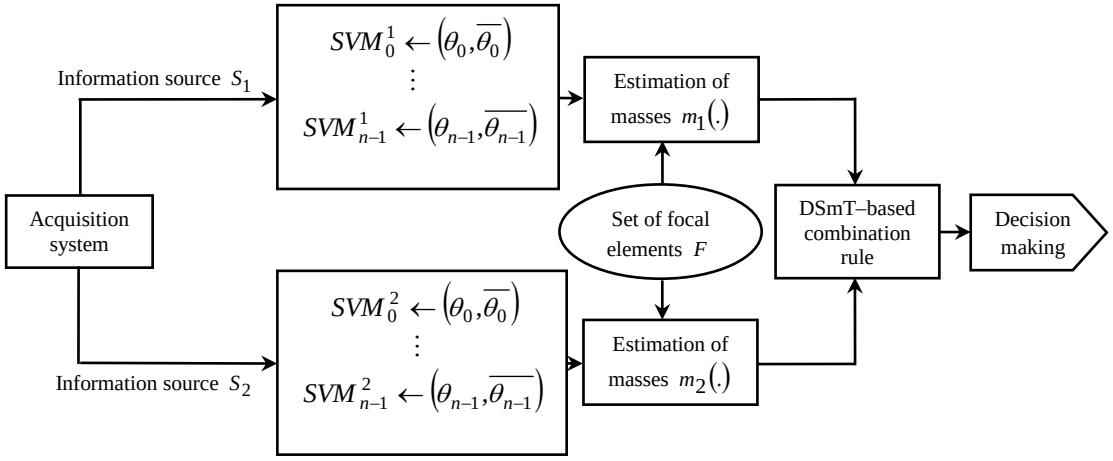


Fig 2. DSmT-based parallel combination for multi-class classification

3.2.1. *Estimation of Masses*

The difficulty of estimating masses is increased if one assigns weights to the composed classes [96]. Therefore, transfer models of the mass function have been proposed whose the aim is to distribute the initial masses on the simple and compound classes associated to each source. Thus, the estimation of masses is performed into two steps: i) assignment of membership degrees for each simple class through a sigmoid transformation proposed by Platt [89], ii) estimation of masses of simple classes and their complementary classes using a supervised model, respectively.

- *Calibration of the SVM outputs:* Although, standard SVM is very discriminative classifier, its output values are not calibrated for appropriately combining two sources of information. Hence, an interesting alternative is proposed in [89] to transform the SVM outputs into posterior probabilities. Thus, given a training set of instance-label pairs $\{(x_k, y_k), k = 1,\ldots,N\}$, where $x_k \in \mathrm{R}^p$ and $y_k \in \{-1,+1\}$, the unthresholded output of an SVM is a distance measure between a test pattern and the decision boundary as given in (5). Furthermore, there is no clear relationship with the posterior class probability $P(y = +1|x)$ that the pattern x belongs to the class $y = +1$. A possible estimation for this

probability can be obtained [89] by modeling the distributions $P(f|y=+1)$ and $P(f|y=-1)$ of the SVM output $f(x)$ using Gaussian distribution of equal variance and then compute the probability of the class given the output by using Bayes' rule. This yields a sigmoid allowing to estimate probabilities:

$$\hat{P}(y=+1|x)=\frac{1}{1+\exp(A\times f(x)+B)} \tag{7}$$

Parameters A and B are tuned by minimizing the negative log-likelihood of the training data:

$$-\sum_{k=1}^{N} t_k \log(Q_k)+(1-t_k)\log(1-Q_k) \tag{8}$$

where $Q_k=\hat{P}(y_k=1|x)$ and $t_k=\frac{y_k+1}{2}$ denotes the probability target.

- *Supervised Model:* Denoting $m_1(.)$ and $m_2(.)$ the gbba provided by two distinct information sources S_1 (First descriptor) and S_2 (Second descriptor), F is the set of focal elements for each source, such that $F=\{\theta_0,\theta_1,\ldots,\theta_{n-1},\overline{\theta_0},\overline{\theta_1},\ldots,\overline{\theta_{n-1}}\}$, the classes θ_i are separable (One relatively to its complementary class $\overline{\theta_i}$) using the SVM-OAA multi-class implementation corresponding to different singletons of the patterns assumed to be known. Therefore, each compound element $A_i \notin F$ has a mass m_1 equal to zero, on the other hand, the mass of the complementary element $\overline{\theta_i}=\bigcup_{\substack{0\le j\le n-1\\ j\neq i}}\theta_j$ is different from zero, which represents the mass of the partial ignorance. The same reasoning is applied to the classes issued from the second source S_2 and $m_2(.)$. Hence, both gbba $m_1(.)$ and $m_2(.)$ are given as follows:

$$m_b(\theta_i)=\frac{\hat{P}_b(\theta_i|x)}{Z_b},\ \forall\theta_i\in F \tag{9}$$

$$m_b(\overline{\theta}_i)=\frac{\sum_{\substack{j=0\\ j\neq i}}^{n-1}\hat{P}_b(\theta_j|x)}{Z_b},\ \forall\overline{\theta}_i\in F \tag{10}$$

$$m_b(A_i)=0,\ \forall A_i\in\Phi=D^{\Theta}\setminus F \tag{11}$$

where $Z_b = \sum_{j=0}^{n-1} \hat{P}_b(\theta_j|x)$ represent the normalization factors introduced in the axiomatic approach in order to respect the mass definition, $\hat{P}_b$ are the posterior probabilities issued from the first source $(b=1)$ and the second source $(b=2)$, respectively. They are given for a test pattern x as follows:

$$\hat{P}_b(\theta_i|x) = \frac{1}{1+\exp(A_{ib} \times f_{ib}(x) + B_{ib})}. \quad (12)$$

A_{ib} and B_{ib} are the parameters of the sigmoid function tuned by minimizing the negative log-likelihood during training for each class of patterns θ_i, and $f_{ib}(x)$ is the i-th output of binary SVM classifier SVM_i^b issued from the source S_b, such that $i = 0,1,\ldots,n-1$ and $b \in \{1,2\}$.

In summary, the masses of all elements $A_j \in D^\Theta$ allocated by each information source S_b $(b=1,2)$ are obtained according the following steps:

1. Define a frame of discernment $\Theta = \{\theta_1,\theta_2,\ldots,\theta_n\}$.
2. Classify a pattern x through the SVM-OAA implementation.
3. Transform each SVM output to the posteriori probability using Eq. (12).
4. Compute the masses associated to each class and its complementary using Eq. (9) and Eq. (10), respectively.

3.2.2. Combination of masses

In order to manage the conflict generated from the two information sources S_1 and S_2 (i.e. both SVM classifications), the combined masses are computed as follows:

$$m_c = m_1 \oplus m_2 \quad (13)$$

where $\oplus$ defines the PCR6 combination rule as given in (1). Hence, in the context of some application of pattern recognition area, such as handwritten digit recognition, we take as constraints the propositions ($\theta_i \cap \theta_j = \emptyset$, $\forall \theta_i,\theta_j \in \Theta$), such that $i \neq j$, which allow separating between each two classes belonging to Θ.

Therefore, the hyper power set D^Θ is reduced to the set F as $F = \{\theta_0,\theta_1,\ldots,\theta_{n-1},\overline{\theta_0},\overline{\theta_1},\ldots,\overline{\theta_{n-1}}\}$, which defines a particular case of the Shafer's model. Thus, the conflict K_c $(\in [0,1])$ measured between two sources is defined as:

$$K_c = \sum_{\substack{A_k, A_l \in F \\ A_k \cap A_l \in \Phi}} m_1(A_k) \times m_2(A_l) \quad (14)$$

where $\Phi = D^{\Theta} \setminus F$ is the set of all relatively and absolutely empty elements, $m_1(.)$ and $m_2(.)$ represent the corresponding generalized basic belief assignments provided by two information sources S_1 and S_2, respectively.

3.2.3. Decision rule

A membership decision of a pattern to one of the simple classes of Θ is performed using the statistical classification technique. First, the combined beliefs are converted into probability measure using a new probabilistic transformation, called Dezert-Smarandache probability (DSmP), that maps a belief measure to a subjective probability measure [14] defined as:

$$DSmP_{\varepsilon}(\theta_i) = m_c(\theta_i) + \left(m_c(\theta_i) + \varepsilon\right) \sum_{\substack{A_j \in 2^{\Theta} \\ A_j \supset \theta_i \\ C_M(A_j) \geq 2}} \frac{m_c(A_j)}{\sum_{\substack{A_k \in 2^{\Theta} \\ A_k \subset A_j \\ C_M(A_k)=1}} m_c(A_k) + \varepsilon C_M(A_j)} \tag{15}$$

where $i = \{0,1,\ldots,9\}$, $\varepsilon \geq 0$ is a tuning parameter, M is the Shafer's model for Θ, and $C_M(A_k)$ denotes the DSm cardinal of A_k [12]. Therefore, the maximum likelihood (ML) test is used for decision making as follows:

$$x \in \theta_i \text{ if } DSmP_{\varepsilon}(\theta_i) = \max\left\{DSmP_{\varepsilon}(\theta_j), 0 \leq j \leq 9\right\} \tag{16}$$

where x is the pattern test characterized by both descriptors, which are used during the feature generation step, and ε is fixed to 0.001 in the decision measure given by (15).

4. Experimental results

4.1. Database description and performance evaluation

For evaluating the effective use of the DSmT for multi-class classification, we consider a case study conducted on the handwriting digit recognition application. For this, we select a well-known US Postal Service (USPS) database that contains normalized grey-level handwritten digit images of 10 numeral classes, extracted from US postal envelopes. All images are segmented and normalized to a size of 16×16 pixels. There are 7291 training data and 2007 test data where some of them are corrupted and difficult to classify correctly (Fig. 3). The partition of the databse for each class according tranining and testing is reported in table 1.

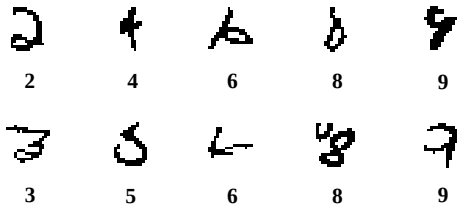


Fig 3. Some samples with their alleged classes from USPS database.

Table 1. Partitioning of the USPS dataset

Classes	0	1	2	3	4	5	6	7	8	9
Training	1194	1005	731	658	652	556	664	645	542	644
Testing	359	264	198	166	200	160	170	147	166	177

For evaluating performances of the handwritten digit classification, a popular error is considered, which is the *Error Rate per Class (ERC)* and *Mean Error Rate (MER)* for all classes. Both errors are expressed in %.

3.2. Pre-processing

The acquired image of isolated digit should be processed to facilitate the feature generation. In our case, the pre-processing module includes a binarization step using the method of Otsu [97], which eliminates the homogeneous background of the isolated digit and keeps the foreground information. Thus, we use the processed digit without unifying size image for recognition process.

3.3. Feature Generation

The objective of the feature generation step is to underline the relevant information that initially exists in the raw data. Thus, an appropriate choice of the descriptor improves significantly the accuracy of the recognition system. In this study, we use a collection of popular feature generation methods, which can be categorized into background features [98,99], foreground features [98,99], geometric features [2], and uniform grid features [100,101].

3.4. Validation of SVM Models

The SVM model is produced for each class according the used descriptor. Hence, the training dataset is partitioned into two equal subsets of samples, which are used for training and validating each binary SVM, respectively. Thus, the validation phase allows finding the optimal hyperparameters for the ten SVM models. In our case, the RBF kernel is selected for the experiments. Furthermore, both the regularization and RBF kernel parameters (C,σ) of each SVM are tuned experimentally during the training phase in such way that the misclassification error of data in the training subset is zero and the validation test gives a minimal error during validation phase for each SVM separating between a simple class and its complementary class.

Table 2 shows an example of the optimal parameters, which are obtained during both training and validation phases by using the UG-SVMs classifier. The parameters n and m define the number of the lines (vertical regions) and columns (horizontal regions) of the grid, respectively, which have been optimized during the validation phase for each SVM model. Therefore, these all parameters are used afterwards during the testing phase. *ERCs* and *ERCc* are the Error Rates per Class for simple and complementary classes, respectively. As we can see, the choice of the optimal size of the uniform grid and hyperparameters of each SVM should be tuned carefully in order to produce a reduced error.

Table 2. Optimal parameters of the UG-SVMs classifier

Parameters	SVM Classifier									
	0	1	2	3	4	5	6	7	8	9
n	7	2	8	5	4	7	7	8	8	7
m	5	3	3	6	12	5	8	6	6	10
σ	3.5	1	3.5	4	3	3.5	4	3.5	5	4.5
C	5	3	4	5	4	4	2	4	3	5
ERCs (%)	2.0	1.0	4.6	5.7	15.6	10.0	2.7	5.5	11.8	4.0
ERCc (%)	0.6	1.1	0.4	0.3	0.1	0.3	0.1	0.1	0.3	0.4

3.5. Quantitative results and discussion

The testing phase is performed using all samples from the test dataset. Hence, the performance of the handwritten digit recognition classification is evaluated on an appropriate choice of descriptors using the SVM classifiers and then we evaluate the combination of the SVMs classifiers within DSmT framework.

3.5.1. Comparative analysis of features

The choice of the complementary features is an important step to ensure efficiently the combination. Indeed, the DSmT-based combination allows offering an accurate performance when the selected features are complementary. Hence, we propose in this section the performance of features in order to select the best ones for combining through the DSmT. For this, we evaluate each SVM-OAA implementation using Foreground Features (FF), Background Features (BF), Geometric Features (GF), Uniform Grid Features (UGF), and the descriptors deduced from a concatenation between at least two simple descriptors such as (BF,FF), (BF,FF,GF) and (UGF,BF,FF,GF). Indeed, the experiments have shown that the appropriate choice of both descriptors and concatenation order to represent each digit class in the feature generation step provides an interesting error reduction. In table 3, FF and UGF-based descriptors using SVM classifiers are evaluated. When concatenating background and foreground (BF,FF)-features, we observe a significant reduction of the MER. Indeed, an error rate reduction of 6.71% is obtained when concatenating BF and FF, respectively. Furthermore, an error rate reduction of 1.5% is obtained when concatenating BF, FF and GF, respectively. This proves that BF, FF and GF are complementary and are more suitable for concatenation. In contrast, when concatenating UGF with BF, FF and GF, the MER is increased to 2.73% comparatively to UGF. This proves that the concatenation does not always allow improving the performance of the classification. Thus, we expect that the UGF and (BF,FF,GF) descriptors are more suitable for combining through the DSmT.

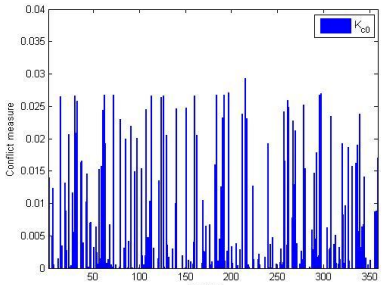
Table 3. Mean error rates of the SVM classifiers using different methods of feature generation

Descriptor	MER (%)
(a) FF	18.87
(b) (BF,FF)	12.16
(c) (BF,FF,GF)	10.66
(d) UGF	6.98
(e) (UGF,BF,FF,GF)	9.71

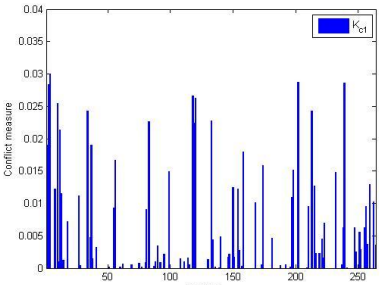
3.5.2. Performance evaluation of the proposed combination framework

In these experiments, we evaluate a handwritten digit recognition classification based on a combination of SVM classifiers through DSmT. The proposed combination framework allows exploiting the redundant and complementary nature of the (BF,FF,GF) and UGF-based descriptors and manage the conflict provided from the outputs of SVM classifiers.

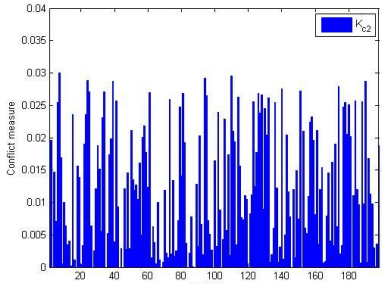
Decision making will be only done on the simple classes belonging to the frame of discernment. Hence, we consider in both combination process and calculation of the decision measures the masses associated to all classes representing the partial ignorance $\overline{\theta_i} = \bigcup_{\substack{0 \le j \le n-1 \\ j \neq i}} \theta_j$ and $\overline{\theta_i} \cap \overline{\theta_j}$ such that $i \neq j$. Thus, in order to appreciate the advantage of combining two sources of information through the DSmT-based algorithm, Figure 4 shows values of the distribution of the conflict measured for each test sample between both SVM-OAA implementations using (BF,FF,GF) and UGF-based descriptors for the 10 digit classes $(\theta_i, i = 0,1,\ldots,9)$, respectively. Table 4 reports the minimal and maximal values of the conflict $(K_{ci}, i = 0,1,\ldots,9)$ generated through the supervised model, which represent the mass assigned to the empty set, after combination process. As we can see, the conflict is maximal for the digit 4 while it is minimal for the digit 9.



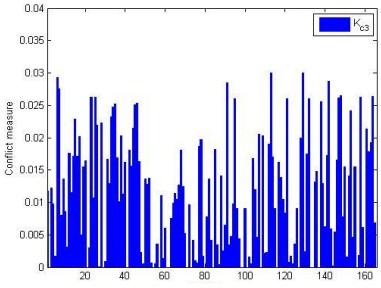
(a) Measured conflict for the digits belonging to θ_0



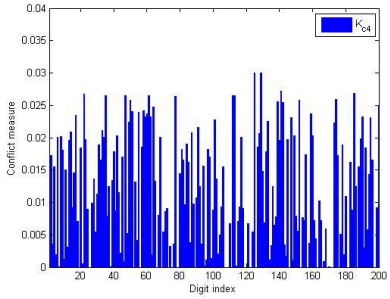
(b) Measured conflict for the digits belonging to θ_1



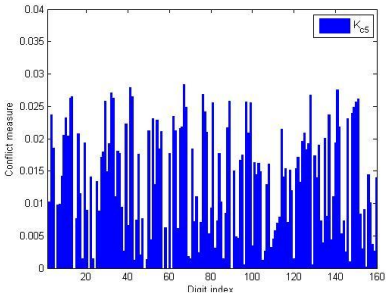
(c) Measured conflict for the digits belonging to θ_2



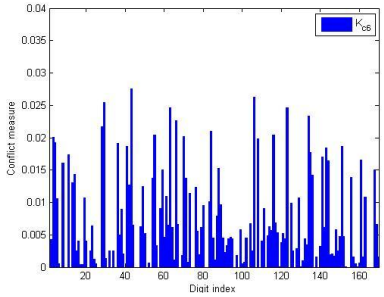
(d) Measured conflict for the digits belonging to θ_3



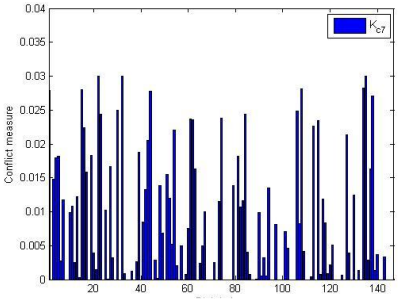
(e) Measured conflict for the digits belonging to θ_4



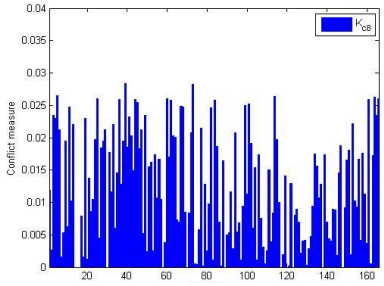
(f) Measured conflict for the digits belonging to θ_5



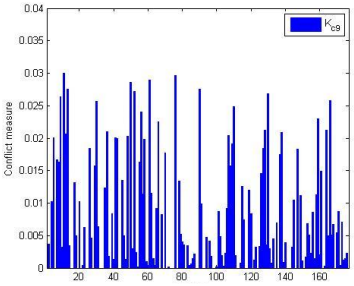
(g) Measured conflict for the digits belonging to θ_6



(h) Measured conflict for the digits belonging to θ_7



(i) Measured conflict for the digits belonging to θ_8



(j) Measured conflict for the digits belonging to θ_9

Fig 4. Conflict between both SVMs classifiers using (BF,FF,GF) and UGF-based descriptors for the ten digit classes $(\theta_i, i = 0,1,\ldots,9)$, respectively.

Table 4. Ranges of conflict variations measured between both SVM-OAA implementations using (BF,FF,GF) and UGF-based descriptors

Class	Minimal conflict (10^{-5})	Maximal conflict (10^{-2})
0	2.149309	2.9933
1	6.999035	2.9964
2	2.747717	2.9992
3	2.936855	2.9994
4	0.494599	3.0000
5	1.868961	2.9970
6	2.537015	2.9887
7	2.826402	2.9983
8	1.485899	2.9910
9	0.276778	2.9999

For an objective evaluation, Table 5 shows ERC and MER produced from three SVM-OAA implementations using UGF, (BF,FF,GF), the descriptor resulting from a concatenation of both UGF and (BF,FF,GF) (i.e. combination at features level) and finally the PCR6 combination rule (i.e. combination at measure level) performed on (BF,FF,GF) and UGF based descriptors, respectively.

Table 5. Error rates of the proposed framework with PCR6 combination rule using (BF,FF,GF) and UGF descriptors

	Descriptor		Concatenation	Combination rule
ERC (%)	(BF,FF,GF)	UGF	(UGF,BF,FF,GF)	PCR6
0	6.69	**1.95**	9.75	**1.95**
1	4.55	3.79	3.79	**3.03**
2	12.63	8.08	**3.54**	6.06
3	17.47	**10.84**	18.67	**10.84**
4	20.00	11.50	19.50	**9.00**
5	16.87	10.00	10.62	**7.50**
6	**2.94**	5.29	4.71	3.53
7	8.84	8.16	8.84	**4.76**
8	12.05	10.84	10.24	**6.63**
9	10.73	6.21	10.17	**5.65**
MER (%)	**10.66**	**6.98**	**9.71**	**5.43**

Overall, the proposed framework using PCR6 combination rule is more suitable than individual SVM-OAA implementations since it provides a MER of 5.43% comparatively to the concatenation which provides a MER of 9.71%. However, when

inspecting carefully each class, we can note that the PCR6 combination rule allows keeping or reducing in the most cases the ERC except for the samples belonging to classes θ_2 and θ_6 .This bad performance is due to the wrong characterization of both UG and (BF,FF,GF)-based descriptors. In other words, the PCR6 combination is not reliable when the complementary information provided from both descriptors is wrongly preserved.

Thus, PCR6 combination rule allows managing correctly the conflict generated from SVM-OAA implementations, even when they provide very small values of the conflict (see Table 4) specifically in the case of samples belonging to θ_8 . Thus, the DSmT is more appropriate to solve the problem for handwritten digit recognition. Indeed, the PCR6 combination rule allows an efficient redistribution of the partial conflicting mass only to the elements involved in the partial conflict. After redistribution, the combined mass is transformed into the DSm probability and the maximum likelihood (ML) test is used for decision making. Finally, the proposed algorithm in DSmT framework is the most stable across all experiments whereas recognition accuracies pertaining to both individual SVM classifiers vary significantly.

4. Conclusion and future work

In this paper, we proposed an effective use of the DSmT for multi-class classification using conjointly the SVM-OAA implementation and a supervised model. Exclusive constraints are introduced through a direct estimation technique to compute the belief assignments and reduce the number of focal elements. Therefore, the proposed framework allows reducing drastically the computational complexity of the combination process for the multi-class classification. A case study conducted on the handwritten digit recognition shows that the proposed supervised model with PCR6 rule yields the best performance comparatively to SVM multi-classifications even when they provide uncalibrated outputs. In continuation to the present work, the next objectives consist to adapt the use of one-class classifiers instead of the OAA implementation of SVM in order to obtain a fixed number of focal elements within DSmT combination process. This will allow us to have a feasible computational complexity independently of the number of combined sources and the size of the discernment space.

References

[1] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification. 2nd edition, Wiley Interscience, New York, 2001.

[2] M. Cheriet, N. Kharma, C.-L. Liu, C.Y. Suen, Character Recognition Systems: A Guide for Students and Practitioner, John Wiley & Sons, 2007.

[3] A. Rahman and M.C. Fairhurst, Multiple classifier decision combination strategies for character recognition: A review, International Journal on Document Analysis and Recognition 5 (4) (2003) 166-194.

[4] L. Xu, A. Krzyzak, C. Suen, Methods of combining multiple classifiers and their applications to handwriting recognition, IEEE Transactions on Systems, Man, and Cybernetics 22 (3) (1992) 418-435.

[5] A. Jain, R. Duin, J. Mao, Statistical pattern recognition: a review, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22 (1) (2000) 4-37.

[6] D. Ruta and B. Gabrys, An overview of classifier fusion methods, Computing and Information Systems 7 (1) 2000 1-10.

[7] N. Abbas, Y. Chibani, H. Nemmour, Handwritten Digit Recognition Based On a DSmT-SVM Parallel Combination, in: Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2012, pp. 241-246.

[8] L.A. Zadeh, Fuzzy algorithm. Information and Control 12 (1968) 94-102.

[9] A.P. Dmpster, Upper and lower probabilities induced by a multivalued maping, Annals of Mathematical Statistics 38 (2) (1967) 325-339.

[10] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[11] D. Dubois, H. Prade, Representation and combination of uncertainty with belief functions and possibility measures, Computational Intelligence 4 (1988) 244-264.

[12] F. Smarandache, J. Dezert, Advances and Applications of DSmT for Information Fusion, American Research Press, 2004.

[13] F. Smarandache and J. Dezert, Advances and Applications of DSmT for Information Fusion, American Research Press, 2006.

[14] F. Smarandache and J. Dezert, Advances and Applications of DSmT for Information Fusion, American Research Press, 2009.

[15] G. Shafer, Perspectives on the Theory and Practice of Belief Functions, International Journal of Approximate Reasoning 4 (1990) 323-362.

[16] Ph. Smets, Practical uses of belief functions, In K.B. Laskey, H. Prade, editors, Fifteenth Conference on Uncertainty in Artifficial Intelligence 99 (1999) 612-621.

[17] G.M. Provan, The validity of Dempster-Shafer Belief Functions, International Journal of Approximate Reasoning 6 (1992) 389-399.

[18] D. Dubois, H. Prade, Evidence, Knowledge and Belief Functions, International Journal of Approximate Reasoning, 6 (1992) 295-319.

[19] E.H. Ruspini, D.J. Lowrance, T.M. Start, Understanding Evidential Reasoning, International Journal of Approximate Reasoning 6 (1992) 401-424.

[20] S. Corgne, L. Hubert-Moy, J. Dezert, G. Mercier, Land cover change prediction with a new theory of plausible and paradoxical reasoning, in: Proceedings of the 6th International Conference on Information Fusion, 2003, pp. 1141-1148.

[21] P. Maupin and A.-L. Jousselme, "Vagueness, a multifacet concept - a case study on Ambrosia artemisiifolia predictive cartography," in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), vol. 1, 2004, pp. 20-24.

[22] A. Elhassouny, S. Idbraim, A. Bekkari, D. Mammass, D. Ducrot, Change Detection by Fusion/Contextual Classification based on a Hybrid DSmT Model and ICM with Constraints, International Journal of Computer Applications 35 (8) (2011) 28-40.

[23] Z. Liu, J. Dezert, G. Mercier, Q. Pan, Dynamical Evidential Reasoning For Changes Detections In Remote Sensing Images, IEEE Transactions on Geoscience and Remote Sensing 50 (5) (2012) 1955-1967.

[24] B. Pannetier, J. Dezert, E. Pollard, Improvement of Multiple Ground Targets Tracking with GMTI Sensor and Fusion of Identification Attributes, Aerospace Conference, 2008, pp. 1-13.

[25] B. Pannetier, J. Dezert, GMTI and IMINT data fusion for multiple target tracking and classification, in: Proceedings of the 12th International Conference on Information Fusion, 2009, pp. 203-210.

[26] P. Kechichian, B. Champagne, An improved partial Haar dual adaptive filter for rapid identification of a sparse echo channel, Signal Processing 89 (5) (2009) 710-723.

[27] Y. Sun, L. Bentabet, A particle filtering and DSmT Based Approach for Conflict Resolving in case of Target Tracking with multiple cues, Journal of Mathematical Imaging and Vision 36 (2) (2010) 159-167.

[28] J. Dezert, B. Pannetier, A PCR-BIMM filter for maneuvering target tracking, in: Proceedings of the 13th International Conference on Information Fusion, 2010.

[29] B. Pannetier, J. Dezert, Extended and multiple target tracking: Evaluation of an hybridization solution, in: Proceedings of the 14th International Conference on Information Fusion, 2011.

[30] R. Singh, M. Vatsa, A. Noore, Integrated Multilevel Image Fusion and Match Score Fusion of Visible and Infrared Face Images for Robust Face Recognition, Pattern Recognition - Special Issue on Multimodal Biometrics 41 (3) (2008) 880-893.

[31] M. Vatsa, R. Singh, A. Noore, Unification of Evidence Theoretic Fusion Algorithms: A Case Study in Level-2 and Level-3 Fingerprint Features, IEEE Transaction on Systems, Man, and Cybernetics - A 29 (1) (2009).

[32] M. Vatsa, R. Singh, A. Noore, M. Houck, Quality-Augmented Fusion of Level-2 and Level-3 Fingerprint Information using DSm Theory, International Journal of Approximate Reasoning 50 (1) (2009).

[33] M. Vatsa, R. Singh, A. Noore, A. Ross, On the Dynamic Selection in Biometric Fusion Algorithms, IEEE Transaction on Information Forensics and Security 5 (3) (2010) 470-479.

[34] E. Garcia, L. Altamirano, Multiple Cameras Fusion Based on DSmT for Tracking Objects on Ground Plane, in: Proceedings of the 11th International Conference on Information Fusion, 2008.

[35] M. Khodabandeh, A. Mohammad-Shahri, Data Fusion of Cameras' Images for Localization of an Object; DSmT-based Approach, in: Proceedings of the 1st International Symposium on Computing in Science & Engineering, 2010.

[36] J. Dezert, Z. Liu, G. Mercier, Edge Detection in Color Images Based on DSmT, in: Proceedings of the 14th International Conference on Information Fusion, 2011.

[37] X. Huang, X. Li, M. Wang, J. Dezert, A FUSION MACHINE BASED ON DSmT AND PCR5 FOR ROBOT'S MAP RECONSTRUCTION, International Journal of Information Acquisition (IJIA) 3 (3) (2006) 201-211.

[38] X. Li, X. Huang, M. Wang, Robot Map Building from Sonar Sensors and DSmT, International Journal of Information & Security 20 (2006) 104-121.

[39] X. Li, X. Huang, M. Wang, Sonar Grid Map Building of Mobile Robots Based on DSmT, Information Technology Journal 5 (2) (2006) 267-272.

[40] X. Li, X. Huang, J. Dezert, L. Duan, M. Wang, A SUCCESSFUL APPLICATION OF DSmT IN SONAR GRID MAP BUILDING AND COMPARISON WITH DST-BASED APPROACH, International Journal of Innovative Computing, Information and Control (ICIC) 3 (3) (2007) 539-549.

[41] P. Li, X. Huang, M. Wang, X. Zeng, Multiple Mobile Robots Map Building Based on DSmT, IEEE International Conference on Robotics, Automation and Mechatronics (RAM) (2008) 509-514.

[42] X. Huang, P. Li, M. Wang, Evidence Reasoning Machine based on DSmT for mobile robot mapping in unknown dynamic environment, in: Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO), 2009, pp. 753-758.

[43] N. Abbas, Y. Chibani, SVM-DSmT combination for the simultaneous verification of off-line and on-line handwritten signatures, International Journal of Computational Intelligence and Applications (IJCIA) 11 (3) 2012.

[44] N. Abbas, Y. Chibani, SVM-DSmT Combination for Off-Line Signature Verification, in: Proceedings of the IEEE International Conference on Computer, Information and Telecommunication Systems (CITS), 2012, pp. 1-5.

[45] M. Tombak, A. Isotamm, T. Tamme, On logical method for counting Dedekind numbers, Lecture Notes in Computer Science, vol. 2138, Springer-Verlag, 2001, pp. 424-427. <www.cs.ut.ee/people/m_tombak/publ.html>

[46] R. Dedekink, Über Zerlegungen von Zahlen durch ihre grössten gemeinsammen Teiler, in: Gesammelte Werke, Bd. 1, 1897, pp. 103-148.

[47] L. Comtet, Sperner Systems, sec. 7.2 in Advanced Combinatorics: The Art of Finite and Infinite Expansion, D. Reidel Publ. Co. (1974) 271-273.

[48] J. Dezert, F. Smarandache, The generation of the hyper-power sets, Chap. 2, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2004, pp. 37-48.

[49] J. Dezert, F. Smarandache, Partial ordering on hyper-power sets, Chap. 3, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2004, pp. 49-60.

[50] R. Kennes, Computational Aspect of the Möbius Transformation of Graphs, IEEE Transaction on Systems, Man, and Cybernetics – Part A: Systems and Humans 22 (2) (1992) 201-223.

[51] Ph. Smets, The application of matrix calculus for belief functions, International Journal of Approximate Reasoning 31 (2002) 1-30.

[52] T. Denoeux, Inner and outer approximation of belief structures using a hierarchical clustering approach, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 9 (4) (2001) 437-460.

[53] P. Djiknavorian, D. Grenier, Reducing DSmT hybrid rule complexity through optimization of the calculation algorithm, Chap. 15, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2006, pp. 345-440.

[54] A. Martin, C. Osswald, A new generalization of the proportional conflict redistribution rule stable in terms of decision, Chap. 2, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2006, pp. 69-88.

[55] A. Martin, Implementing general belief function framework with a practical codification for low complexity, Chap. 7, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2009, pp. 217-273.

[56] J. Dezert, F. Smarandache, Combination of beliefs on hybrid DSm models, Chap. 4, in: Advances and Application of DSmT for Information Fusion, American Research Press, 2004, pp. 61-103.

[57] X. Li, J. Dezert, F. Smarandache, X. Huang, Evidence Supporting Measure of Similarity for Reducing the Complexity in Information Fusion, Information sciences 181 (10) (2011) 1818–1835.

[58] P.S. Deng, H.Y.M. Liao, C.-W. Ho, H.-R. Tyan, Wavelet based off-line handwritten signature verification, Computer Vision and Image Understanding 76 (3) (1999) 173–190.

[59] B. Fang, C.H. Leung, Y.Y. Tang, K.W. Tse, P.C.K. Kwok, Y.K. Wong, Off-line signature verification by the tracking of feature and stroke positions, Pattern Recognition 36 (2003) 91–101.

[60] J.K. Guo, D. Doermann, A. Rosenfeld, Forgery detection by local correspondence, International Journal of Pattern Recognition and Artificial Intelligence 15 (4) (2001) 579–641.

[61] B. Fang, Y.Y. Wang, C.H. Leung, K.W. Tse, Y.Y. Tang, P.C.K. Kwok, Y. K. Wong, Offline signature verification by the analysis of cursive strokes, International Journal of Pattern Recognition and Artificial Intelligence 15 (4) (2001) 659–673.

[62] R. Sabourin, G. Genest, F. Prêteux, Off-line signature verification by local granulometric size distributions, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (9) (1997) 976–988.

[63] E.J.R. Justino, F. Bortolozzi, R. Sabourin, A comparison of SVM and HMM classifiers in the off-line signature verification, Pattern Recognition Letters 26 (2005) 1377-1385.

[64] E.J.R. Justino, F. Bortolozzi, R. Sabourin, Off-line Signature Verification Using HMM for Random, Simple and Skilled Forgeries, in: Proceedings of the International Conference on Document Analysis and Recognition (ICDAR), vol. 1, 2001, pp. 105-110.

[65] J. Coetzer, B.M. Herbst, J.A. de Preez, Offline Signature Verification Using the Discrete Radon Transform and a Hidden Markov Model, EURASIP Journal on Applied Signal Processing 4 (2004) 559-571.

[66] T. Kaewkongka, K. Chamnongthai, B. Thipakorn, Offline signature recognition using parameterized hough transform, in: Proceedings of the 5th International Symposium on Signal Processing and Its Applications, 1999, pp. 451–454.

[67] C. Quek, R.W. Zhou, Antiforgery: a novel pseudo-outer product based fuzzy neural network driven signature verification system, Pattern Recognition Letters 23 (14) (2002) 1795–1816.

[68] J. Weston, C. Watkins, Support Vector Machines for Multi-Class Pattern Recognition Machines, in: CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, 1998.

[69] J. Weston, C. Watknis, Multi-class support vector machines, Tech. Rep. CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, Egham, TW20 0EX, UK, 1998.

[70] Y. Guermeur, A. Elisseeff, H. Paugam-Moisy, Estimating the sample complexity of a multi-class discriminant model, In: Proceedings of the Industrial Conference on Artificiel Neural Networks, 1999, pp. 310-315.

[71] C. Hsu, C. Lin, A comparaison of methods for muli-class support vector machines, Tech. Rep, Department of Computer Science and Information Engineering, National Taiwan University, 2001.

[72] L. Bottou, C. Cortes, J. Drucker, I. Guyon, L. LeCunn, U. Muller, E. Sackinger, P. Simard, V. Vapnik, Comparaison of Classifier methods: a case study in handwriting digit recognition, In: Proceedings of the International Conference on Pattern Recognition, 1994, pp. 77-87.

[73] H.P. Huang, Y.H. Liu, Fuzzy support vector machines for pattern recognition and data mining, International Journal of Fuzzy Systems 4 (3) (2002) 826-835.

[74] A. Martin, C. Osswald, Toward a combination rule to deal with partial conflict and specificity in belief functions theory, in: Proceedings of the International Conference on Information Fusion, 2007.

[75] A. Aregui, T. Denoeux, Fusion of one-class classifier in the belief function framework, in: Proceedings of the International Conference on Information Fusion, 2007.

[76] B. Quost, T. Denoeux, M. Masson, Pairwise classifier combination using belief functions, Pattern Recognition Letters 28 (5) (2007) 644–653.

[77] B. Quost, T. Denoeux, M. Masson, Pairwise classifier combination in the framework of belief functions, in: Proceedings of the International Conference on Information Fusion, 2005.

[78] B. Quost, T. Denoeux, M. Masson, One-against-all combination in the framework of belief functions, in: Proceedings of the IPMU , vol. 1, 2006, pp. 356-363.

[79] B. Quost, T. Denoeux, M. Masson, Combinaison crédibiliste de classifieurs binaires, Traitement du Signal 24 (2) (2007) 83-101.

[80] Z. Hu, Y. Li, Y. Cai, X. Xu, Method of combining multi-class SVMs using Dempster-Shafer theory and its application, in: Proceedings of the American Control Conference, 2005.

[81] A. Martin, I. Quidu, Decision support with belief functions theory for seabed characterization, in: Proceedings of the International Conference on Information Fusion, 2008.

[82] T. Burger, O. Aran, Modeling hesitation and conflict: a belief-based approach for multi-class problems, in: Proceedings of the 5th International Conference on Machine Learning and Applications-ICMLA, 2006, pp. 95-100.

[83] T. Burger, A. Urankar, O. Aran, L. Akarun, A. Caplier, A Dempster-Shafer theory based combination of classifiers for hand gesture recognition, Computer Vision and Cumputer Graphics: Theory and Applications 21 (2008) 137-150.

[84] H. Laanaya, A. Martin, D. Aboutajdine, A. Khenchaf, Support vector regression of membership functions and belief functions-Application for pattern recognition, Information Fusion 11 (4) (2010) 338–350.

[85] T. Denoeux, Analysis of evidence theoretic decision rules for pattern classification, Pattern recognition 30 (7) (1997) 1095-1107.

[86] Ph. Smets, R. Kennes, The transferable belief model, Artificial Intelligence 66 (1994) 191-234.

[87] D. Dubois, H. Prade, Fuzzy Sets-A Convenient Fiction for Modeling Vagueness and Possibility, IEEE Transactions on Fuzzy Systems 2 (1) (1994).

[88] A. Appriou, Multisensor signal processing in the framework of the theory of evidence, NATO/RTO, Application of Mathematical Signal Processing Techniques to Mission Systems, 1999.

[89] J.C. Platt, Probabilities for SV Machines, Advances in Large Margin Classifiers, MIT Press, 1999, pp. 61-74.

[90] I. Bloch. Fusion d'informations en traitement du signal et des images. IC2. Hermès Science, 2003.

[91] S. French, Group Consensus Probability Distributions: A Critical Survey, In J. Bernardo et al., editor, Bayesian Statistics, Elsevier, 1985, pp. 183-201.

[92] R. Cooke, Uncertainty in Risk Assessment: A Probabilist's Manifesto, Reliability in Engineering and Systems Safety 23 (1988) 277-283.

[93] R. Cooke. Experts in Uncertainty. Oxford University Press, 1991.

[94] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, 1995.

[95] D. Cortes, V. Vapnik, Support vector networks, Machine Learning 20 (1995) 273-297.

[96] J.D. Lowrance, T.M. Strat, L.P. Wesley, T.D. Garvey, E.H. Ruspini, D.E. Wilkins, The Theory, Implementation and Practice of Evidential Reasoning, SRI project 5701 final report, SRI, Palo Alto, 1991.

[97] N. Otsu, A Threshold Selection Method from Gray-Level Histograms, IEEE Transactions on Systems, Man, and Cybernetics 9 (1979) 62-66.

[98] A. Britto, R. Sabourin, F. Bortolozzi, C. Suen, Foreground and Background Information in an HMM-based Method for Recognition of Isolated Characters and Numeral Strings, in: Proceedings of the International Workshop on Frontiers in Handwriting Recognition, 2004, pp. 371-376.

[99] P.R. Cavalin, A. Britto, F. Bortolozzi, R. Sabourin, L. Oliveira, An implicit segmentation-based method for recognition of handwritten strings of characters, ACM symposium on Applied computing, 2006, pp. 836-840.

[100] J. Favata, G. Srikantan, A Multiple Feature/Resolution Approach To Handprinted Digit and Character Recognition, Proceedings of the International journal of imaging systems and technology 7 (4) (1996) 304–311.

[101] N. Abbas, Y. Chibani, Combination of Off-Line and On-Line Signature Verification Systems Based on SVM and DST, in: The 11th International Conference on Intelligent Systems Design and Applications, 2011, pp. 855 – 860.

Intelligent Alarm Classification Based on DSmT

Albena Tchamova
Jean Dezert

Originally published as Tchamova A., Dezert J., *Intelligent Alarm Prioritization based on DSmT*, IEEE Intelligent Systems IS'2012, Sofia, Bulgaria, Sept. 6-8, 2012, and reprinted with permission.

***Abstract*—In this paper the critical issue of alarms' classification and prioritization (in terms of degree of danger) is considered and realized on the base of Proportional Conflict Redistribution rule no.5, defined in Dezert-Smarandache Theory of plausible and paradoxical reasoning. The results obtained show the strong ability of this rule to take care in a coherent and stable way for the evolution of all possible degrees of danger, relating to a set of a priori defined, out of the ordinary dangerous directions. A comparison with Dempster's rule performance is also provided. Dempster's rule shows weakness in resolving the cases examined. In Emergency case Dempster's rule does not respond to the level of conflicts between sound sources, leading that way to ungrounded decisions. In case of lowest danger's priority (perturbed Warning mode), Dempster's rule could cause a false alarm and can deflect the attention from the existing real dangerous source by assigning a wrong steering direction to the surveillance camera.**

***Keywords*—Alarm classification; DSmT; DST; data fusion.**

I. Introduction

The alarms classification and prioritization is a very challenging and difficult task. The encountered overflowing amount of alarms could become a serious source of confusion especially in dangerous cases, when one needs to take a proper immediate response. The problem is really critical, because the information available for performing alarms processing is uncertain, imprecise, even conflicting. There are cases, when some of the alarms generated could be incorrectly interpreted as false, increasing the chance to be ignored, in case when they are really significant and dangerous. That way the critical delay of the proper response could cause significant damages.

A lot of work was done during the years, because the importance of this problem was recognized since the 1960s, in wide world cases of surveillance: in industry (powerplants, oil refineries), the clinical alarms in medicine, civilian and military monitoring. Nowadays surveillance (military and civilian) and environmental monitoring systems are characterized with a smart operational control, based on the intelligent analysis and interpretation of alarms coming from a variety of sensors installed in the observation area. Many approaches have been adopted and applied, addressing the problem in common. In [1] a generic neuro-expert system architecture for training neural networks in alarm processing is developed, which is satisfactory when the training set covers enough range of scenarios. An expert system with temporal reasoning for alarm processing is proposed in [2]. Fault detection and alarm processing in a loop system using a fault detection system is presented in [3]. In [4] the authors consider a methodology, based on both artificial neural networks and fuzzy logic for alarm identification. The tasks of alarm processing, fault diagnosis and comprehensive validation of protection performance are discussed and resolved in [5] using knowledge-based systems and model-based reasoning approach. In [6] alarm prioritization, using fuzzy logic is developed to prioritize the alarms during alarm floods which would ease the burden of operators with meaningless or false alarms. In case of multiple suspicious signals, generated from a number of sensors in the observed area, the problem of alarm classification requires the most dangerous among them to be correctly recognized, in order to decide properly where the video camera should be oriented. Because of uncertainty and conflicts encountered in signals' data, one needs to process, analyze and interpret correctly in timely manner all suspicious sound signals separately at particular sensor's levels in the observed area. Such kind of conflicts could weaken or even mistake the decision about the degree of danger in a critical situation. That is why a strategy for an intelligent, scan by scan, combination/updating of sounds data generated by each sensor is needed in order to provide the surveillance system with a meaningful output. There are various well known methods for combining information, which could be applied. The most used until now Dempster-Shafer Theory (DST) [9] proposes a suitable mathematical model for uncertainty representation, but its weak point in applications relates to the normalization factor, which yields to non-adequate results when sources to combine are highly conflicting. To overcome such drawback, we apply the Proportional Conflict Redistribution Rule no.5 (PCR5), defined in Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning [7]. It proposes a powerful and efficient way for combining and utilizing all the available information, allowing the possibility for conflicts and paradoxes between the elements of the frame of discernment. A comparison with DST performance based on Dempster's rule of combination[1] is also provided in order to evaluate the ability of DSmT to assure awareness about the alarms' classification and prioritization in case of sound source data discrepancies and to improve decision-making process about the degree of danger. In section II we recall basics of DST and

[1]This rule is also called Dempster-Shafer rule, and denoted DS for short.

Dempster's rule. Basics of PCR5 fusion rule are outlined in section III. Section IV relates to the decision making support used in order to decide which sound source is most dangerous. In section V, we present the problem of alarms classification and examine two solutions to solve it by using PCR5 and Dempster's rule. In section VI, the evaluation and comparative analysis of both solutions are provided on a given simulation scenario, that includes three sensors, generating three types of signals (warning, alarm and emergency). Concluding remarks are given in section VII.

II. Basics of DST

DST [9] proposes a suitable mathematical model for uncertainty representationLet $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ be a frame of discernment of a problem under consideration containing n distinct elements θ_i, $i = 1, \ldots, n$. A basic belief assignment (bba, also called a belief mass function) $m(.) : 2^\Theta \rightarrow [0, 1]$ is a mapping from the power set of Θ (i.e. the set of subsets of Θ), denoted 2^Θ, to $[0, 1]$, that must satisfy the following conditions: 1) $m(\emptyset) = 0$, i.e. the mass of empty set (impossible event) is zero; 2) $\sum_{X\in 2^\Theta} m(X) = 1$, i.e. the mass of belief is normalized to one. $m(X)$ represents the mass of belief exactly committed to X. The vacuous bba characterizing full ignorance is defined by $m_v(.) : 2^\Theta \rightarrow [0; 1]$ such that $m_v(X) = 0$ if $X \neq \Theta$, and $m_v(\Theta) = 1$. From any bba $m(.)$, the belief function $Bel(.)$ and the plausibility function $Pl(.)$ are defined as $\forall X \in 2^\Theta : Bel(X) = \sum_{Y|Y\subseteq X} m(Y)$ and $Pl(X) = \sum_{Y|X\cap Y\neq\emptyset} m(Y)$. $Bel(X)$ and $Pl(X)$ are classically seen as lower and upper bounds of an unknown probability $P(X)$ of X. Dempster-Shafer (DS) rule of combination [9] is a mathematical operation, denoted $\oplus$, which corresponds to the normalized conjunctive fusion rule. Based on Shafer's model of the frame, the combination of two independent and distinct sources of evidences characterized by their bba $m_1(.)$ and $m_2(.)$ and related to the same frame of discernment Θ is defined by $m_{DS}(\emptyset) = 0$, and $\forall X \in 2^\Theta \setminus \{\emptyset\}$ by

$$m_{DS}(X) = [m_1 \oplus m_2](X) = \frac{m_{12}(X)}{1 - K_{12}} \tag{1}$$

where

$$m_{12}(X) \triangleq \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) \tag{2}$$

corresponds to the conjunctive consensus on X between the two sources of evidence. K_{12} is the total *degree of conflict* between the two sources of evidence defined by

$$K_{12} \triangleq m_{12}(\emptyset) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2) \tag{3}$$

DS rule is commutative and associative. The weak point of this rule is its behavior when $K_{12} \rightarrow 1$ because it can generate unexpected (at least very disputable) results [11]. When $K_{12} = m_{12}(\emptyset) = 1$, the two sources are said to be in total conflict and their combination cannot be applied since DS rule is mathematically not defined because of $0/0$ indeterminacy [9].

III. Basics of PCR5 fusion rule

The idea behind the Proportional Conflict Redistribution rule no. 5 (see [7], Vol. 3) is to transfer conflicting masses (total or partial) proportionally to non-empty sets involved in the model according to all integrity constraints. The general principle of PCR rules is then to: *1*) calculate the conjunctive consensus between the sources of evidences; *2*) calculate the total or partial conflicting masses; *3*) redistribute the conflicting mass (total or partial) proportionally on non-empty sets involved in the model according to all integrity constraints. Under Shafer's model assumption of the frame Θ, the PCR5 combination rule for only two sources of information is defined as: $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{Y\in 2^\Theta\setminus\{X\} \\ X\cap Y=\emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)}\right] \tag{4}$$

where $m_{12}(X)$ corresponds to the conjunctive consensus on X between the two sources and where all denominators are different from zero. All sets involved in the formula are in canonical form. All denominators are different from zero. If a denominator is zero, that fraction is discarded. No matter how big or small the conflicting mass is, PCR5 mathematically does a better redistribution of the conflicting mass than DS since PCR5 goes backwards on the tracks of the conjunctive rule and redistributes the partial conflicting masses only to the sets involved in the conflict and proportionally to their masses put in the conflict, considering the conjunctive normal form of the partial conflict. PCR5 is quasi-associative and preserves the neutral impact of the vacuous belief assignment.

IV. Decision-making support

In this work, we assume Shafer's model and we use the classical Pignistic Transformation [7], [10] to take a decision about the mode of danger. The pignistic probability (Pign.Proba), also called the betting probability (BetP) is defined for $\forall A \in 2^\Theta$ by

$$BetP(A) = \sum_{X\in D^\Theta} \frac{|X \cap A|}{|X|} \cdot m(X) \tag{5}$$

where $|X|$ denotes the cardinality of X.

V. Alarms Classification Approach

Our approach for alarms classification assumes all the localized sound sources to be subjects of attention and investigation for being indication of dangerous situations. The specific attributes of input sounds, emitted by each source, are sensor's level processed and evaluated in timely manner for their contribution towards correct alarms' classification (in term of degree of danger). The input sounds attributes generated by each sensor, at each time moment (scan) concern the frequency of intermittence, f_{int} and sound signal duration, T_{sig}. A particular relationship between the specific values of f_{int} and

associated corresponding degree of danger is established, i.e to map input specific sensor level data into the frame of discernments, concerning the level of abstraction *Degree of Danger*= {*Emergency*, *Alarm*, *Warning*}. Then the process consists in temporal sensors' level sound signals' attribute updating on the base of PCR5 fusion rule. Our motivation for attribute fusion is inspired from the necessity to ascertain the degree of danger, associated with all localized sound sources separately, in order to quickly focus on the most dangerous alarm information and to take immediate and correct feedback actions to decide properly where the video camera should be oriented. The applied algorithm considers the following steps:

• We define the frame of expected hypotheses according to the respective degree of danger associated with the attributes's specific values as follows: $\Theta = \{\theta_1 = (E)mergency, \theta_2 = (A)larm, \theta_3 = (W)arning\}$. The hypothesis with a highest priority is *Emergency*, following by *Alarm* and then *Warning*. These hypotheses are exclusive and exhaustive, hence Shafer's model holds and we work on power-set: $2^{\Theta} = \{\emptyset, E, A, W, E \cup A, E \cup W, A \cup W, E \cup A \cup W\}$.

• A rule-base is defined in order to establish the relationships between the sounds' attributes associated with all localized sources and corresponding degrees of danger, in the form:

Rule 1: if *attributes-type 1* then *Emergency*

Rule 2: if *attributes-type 2* then *Alarm*

Rule 3: if *attributes-type 3* then *Warning*

where attributes types 1, 2 and 3 could be specific sounds' attributes values, which are informative enough to be processed and evaluated for their contribution towards correct alarms' classification. In this rule base *attributes-type 1* is a sound's attribute, which is typical for degree of danger *Emergency*, *attributes-type 2* is typical for *Alarm*, *attributes-type 3* for *Warning*. In our case the frequency of intermittencies (if the signal is intermittent) f_{int}, associated with the localized sound sources is utilized. Then the following specific rule-base is used as an input interface to map the sounds' attributes (so called *observations*) obtained from all localized sources into non-Bayesian basic belief assignments $m_{obs}(.)$:

Rule 1: if $f_{int} \to 1Hz$ then $m_{obs}(E) = 0.9$ and $m_{obs}(E \cup A) = 0.1$.

Rule 2: if $f_{int} \to 5Hz$ then $m_{obs}(A) = 0.7$, $m_{obs}(A \cup E) = 0.2$ and $m_{obs}(A \cup W) = 0.1$.

Rule 3: if $f_{int} \to 0Hz$ then $m_{obs}(W) = 0.6$ and $m_{obs}(W \cup A \cup E) = 0.4$.

If the value of the sound attribute received is close to the particular sound signal parameter for Emergency, our bba is constructed in way that it will consider the hypothesis *Emergency* and also the reasonable in this case composite proposition ($E \cup A$), representing a possible partial uncertainty. If the value obtained is close to the particular sound signal parameter for *Alarm*, our bba is constructed in way that it will consider the hypothesis *Alarm* itself and also the reasonable in that case composite propositions $A \cup E$ and $A \cup W$. Assigning a higher mass of belief to $A \cup E$ than to $A \cup W$ is to take care about the possibility for *Emergency* case. If the value obtained is close to the particular sound signal parameter for *Warning*, our bba is constructed in way that it will consider the hypothesis *Warning* and also the composite proposition $E \cup A \cup W$, representing the case of full ignorance, in order to take care about possibility for *Alarm* and especially for *Emergency* case. All the belief masses not already assigned to singletons (E, A or W) are assigned to the reasonable partial uncertainties reflecting the possible noise perturbations in the observed information.

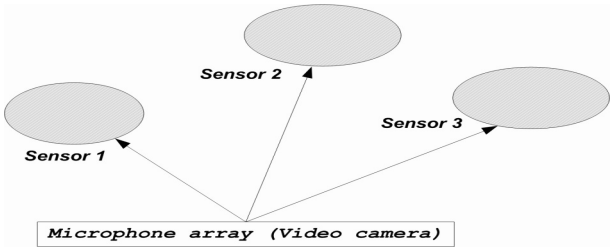


Fig. 1. Scenario.

• At the very first time moment $k = 0$ we start with a priori basic belief assignment (history) set to be a vacuous belief assignment $m_{hist}(E \cup A \cup W) = 1$, since there is no information about the first detected degree of danger according to sound sources.

• Combination of currently received measurement's bba $m_{obs}(.)$ (for each of located sound sources), based on the input interface mapping, with a history's bba, in order to obtain estimated bba relating to the current degree of danger $m(.) = [m_{hist} \oplus m_{obs}](.)$. PCR5 and DS are tested in the process of temporal data fusion to update bba's associated with each sound emitter.

• Flag for an especially high degree of danger has to be taken, when during the a priori defined scanning period, the maximum Pignistic Probability [7] is associated with the hypothesis *Emergency*.

For security purpose, it is very important to keep updating sequentially the estimation one has on the state of the true modes of sound emitters, even if they are in the lowest priority mode (i.e. in warning mode only) in order to prevent unexpected alarm's changes.

VI. Simulation Scenario and Results

In our simulation scenario (Fig. 1) a set of three sensors located at different distances from the microphone array are installed in an observed area for protection purposes, together with a video camera [8]. It is assumed, that sensors are assembled with alarm devices, as follows: Sensor 1 with *Sonitron*, Sensor 2 with *E2S*, and Sensor 3 with *System Sensor* companies alarm devices. In case of alarm events (smoke, flame, intrusion, etc.) the alarm devices emit powerful sound signals with various duration and frequency of intermittence depending on the nature of the event. dangerous signal source. These sensors are used for the purpose of estimation the level of danger/threat for each place where they are located. Data, obtained from each source are processed and analyzed at particular sensor's level independently, in consecutive time

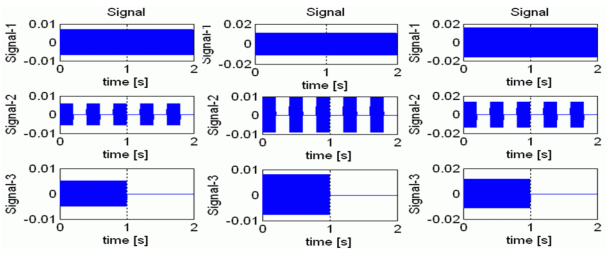


Fig. 2. *Sonitron, E2S, System Sensor* Sound Characteristics.

moments, with regard to all possible degrees of danger: $\theta_1 = (E)mergency$, $\theta_2 = (A)larm$, and $\theta_3 = (W)arning$. Doing this one could find the first suspicious moment, when the situation could become eventually dangerous.

Table 1 Sound signal parameters.

Continuous (*Warning*)	Intermittent-I (*Alarm*)	Intermittent-II (*Emergency*)
$f_{int} = 0$Hz	$f_{int} = 5$Hz	$f_{int} = 1$Hz
$T_{sig} = 10$s	$T_{sig} = 30$s	$T_{sig} = 60$s

The sound signals representing *Warning*, *Alarm* and *Emergency*, emitted from alarm devices, produced by Sonitron, E2S and System Sensor companies used in our simulation (Table 1) are shown on Fig. 2. The first (left) column of Fig. 2 relates to Sonitron, the second column to E2S, and the third (right) column relates to System Sensor devices. The first row of this figure represents the signal 1 for *Warning*, second row represents signal 2, for *Alarm*, and the last third row represents signal 3, for *Emergency* case. The *Alarm* signal is intermittent with a frequency of intermittence $f_{int} = 5Hz$ and a duration $T_{sig} = 30s$, so called type I. The *Emergency* sound signal is intermittent with a frequency of intermittence $f_{int} = 1Hz$ and duration $T_{sig} = 60s$, so called type II. The *Warning* signal is continuous with $f_{int} = 0Hz$ and $T_{sig} = 10s$.

Our simulation scenario considers a true degree of danger associated with the sound sources as follows: Emergency mode for the first sound emitter, Alarm mode for the second, and Warning mode - for the third one. The three sources are processed in parallel and because of possible sound perturbations we assume that possible random changes can be observed over the scans for a given mode. We therefore introduce some switches between the three modes *Emergency*, *Alarm* and *Warning* to simulate what can happen in practice (what we call *ground truth* and displayed with black plots on our next figures 3 and 4. According to this, three main cases are estimated:

- The most interesting for us it is the estimation of danger level by sensor 1, associated with *Emergency* mode. In our simulation, the *The Ground Truth* associated with Sensor 1 considers that during scans 1–3 the observations generated support the *Emergency* mode (the highest level of danger). From scan 4 to scan 6 the observations generated support the *Warning* mode (the lowest level of danger). From scan 7 to scan 30 the observations generated support again *Emergency* mode. Such kind of scenario is important in the real world cases because sources data can be deteriorated by noise perturbations and therefore some possible conflicts arise between observations from scan to scan. We assume that a conflict occurs in sounds data between *Emergency* and *Warning* modes, because it could weaken strongly the decision taken. It could become a reason to ignore the significance of out of ordinary, dangerous situation.

- The second interesting case concerns the estimation of probabilities of modes, associated with the sound emitter 2 working in *Alarm* mode. The *Ground Truth* has been a little bit changed with respect to the ground truth simulated for sensor 1. We assume that during scans 1–3 the observations generated support correctly the *Alarm* mode. From scan 4 to scan 8 the observations generated support the *Emergency* mode because of noise perturbations. From scan 9 to scan 30 the observations generated support again correctly the *Alarm* mode.

- The third interesting case concerns the estimation of the probability of modes, associated with the third emitter working in *Warning* mode. In our simulation of this case, we considers that during scans 1–2 the observations generated support correctly the *Warning* mode. From scan 3 to scan 5 the observations generated support the *Emergency* mode because of some possible noise perturbations. From scan 6 to scan 30 the observations generated support again correctly the *Warning* mode.

As a result of processing and analyzing sounds' data, obtained from the three sources, processed in parallel, one establishes at each scan, for each source the Pignistic probabilities, associated with all the considered modes of danger. The decisions should be governed at the video camera level, taken periodically, depending on: *1)* specificities of the video camera (time needed to steer the video camera toward a localized direction); *2)* time duration needed to analyze correctly and reliably the sequentially gathered information. We choose as a reasonable sampling period for camera decisions $T_{dec} = 20sec$, i.e. at every 10th scan, we should establish the decision about the most probable mode of danger, associated with each sound source, that way to declare directions for steering the video camera. For our scenario, the decisive scans will be 10th, 20th, and 30th. In the next two subsections we analyze the performances of PCR5 and DS to conclude on their ability (or inability) to correctly identify the alarm modes for the prioritization purpose.

A. *PCR5 rule performance for danger level estimation.*

Figure 3 shows the values of Pignistic Probabilities of each mode (*Emergency*, *Alarm*, *Warning*) associated with three sound emitters (1st source in *Emergency* mode, (subplot on the top), 2nd source in *Alarm* mode (subplot in the middle), and 3rd source in *Warning* mode, (subplot in the bottom)) during the all 30 scans. Each source has been perturbed with noises in

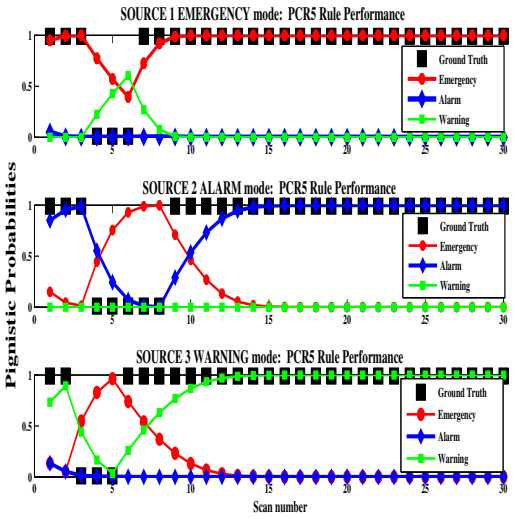


Fig. 3. PCR5 rule Performance for danger level estimation.

accordance with the simulated *Ground Truth*, associated with particular sound source. These probabilities are obtained for each source independently as a result of sequential data fusion of $m_{obs}(.)$ sequence using PCR5 combinational rule. For each source, we analyze the probabilities of its modes obtained with BetP computed from PCR5 rule and the corresponding decisions for steering the camera at scans no. 10, 20, and 30.

Decision taken by PCR5 rule at scan 10*:*

For source 1, associated with *Emergency* mode (Fig. 3, top-subplot), Pign.Proba established by PCR5 at scan 10 are as follows: $BetP(E) = 1.0$, $BetP(A) = 0$, and $BetP(W) = 0$. During the first scans one has $BetP(E) < 1$ because of the impact of the full uncertainty at the beginning. During the transition period between scans 4 and 6 the Pignistic Probability $BetP(E)$ decreases near to 0.4, and in a meantime $BetP(W)$ increases near to 0.6, reflecting that way the new observations supporting the *Warning* mode. After reestablishing the proper sound signal at scan 7, the PCR5 combination rule leads to quick re-estimation of belief masses, assigned to the right *Emergency* mode. One sees clearly the efficiency of PCR5 to detect a mode switch from the sequential fusion of $m_{obs}(.)$. At this processing stage, after decisive 10th scan, PCR5 rule takes a correct, reliable decision that $BetP(E) = 1.0$, assuring that camera will steer at this direction with highest priority.

For source 2, associated with *Alarm* mode, (Fig. 3, middle-subplot), Pign.Proba established by PCR5 are as follows: $BetP(E) = 0.5$, $BetP(A) = 0.5$, and $BetP(W) = 0$. At first scans, $BetP(A) < 1$, because of the full uncertainty at the very first time moment, and then $BetP(A) \to 1$. During the transition time between scans 4 and 8, $BetP(A)$ gradually decreases, while $BetP(E)$ gradually increases. During this period PCR5 rule takes attention according to the mode with the highest priority, i.e. the *Emergency* mode. Starting from scan 9 PCR5 rule reestablishes gradually (and enough quickly after a short delay) the probability mass assigned to *Alarm* mode. At the end of scan 10 PCR5 rule keeps $BetP(A) \approx BetP(E)$, staying cautious about *Emergency*, but this rule is on the way of fully reestablishing the beliefs in the proper *Alarm* mode for this case and to forget the mistaken *Emergency* mode.

For source 3, associated with *Warning* mode, (Fig. 3, subplot in the bottom), Pign.Proba established by PCR5 are as follows: $BetP(E) = 0.2$, $BetP(A) = 0$, and $BetP(W) = 0.8$. Until scan 10, because of the sound attributes measurement conflicts, the PCR5 rule gives some support (non null probability) to *Emergency* mode and also to *Warning* mode. Until scan 10, its behavior is cautious about *Emergency* mode, and during this time period it doesn't establish a hard decision. PCR5 results makes sense, because the decision about *Warning* mode is not decisive/firm.

Decision taken by PCR5 rule at scan 20 *and scan* 30*:*

From scan 15 on, and for all sound sources 1,2 and 3, PCR5 rule estimation is fully adequate and reasonable.

For source 1, associated with *Emergency* mode, one has: $BetP(E) = 1$, $BetP(A) = 0$, and $BetP(W) = 0$.

For source 2, associated with *Alarm* mode: $BetP(E) = 0$, $BetP(A) = 1$, and $BetP(W) = 0$.

For source 3, associated with *Warning* mode: $BetP(E) = 0$, $BetP(A) = 0$, and $BetP(W) = 1$.

These Pign.Proba remain firmly one and the same at scans 20 and 30, associating in stable way the highest priority danger to sound source 1 as expected in such scenario.

B. Dempster's rule performance for danger level estimation.

The corresponding figure 4 shows the values of Pignistic Probabilities of each mode (*Emergency*, *Alarm*, *Warning*) associated with three sound emitters (1st source in *Emergency* mode, (top subplot), 2nd in *Alarm* mode (middle subplot), and 3rd in *Warning* mode (bottom subplot)) during all 30 scans, which are obtained as a result of sequential data fusion of $m_{obs}(.)$ sequence using DS of combination.

Decision taken by Dempster's rule at scan 10*:*

For source 1, associated with *Emergency* mode (Fig. 4, subplot on the top), Pign.Proba established by DS are as follows: $BetP(E) = 1$, $BetP(A) = 0$, and $BetP(W) = 0$. It is obvious, that during the scans 1 and 10 DS is unable to respond to the new observations, arriving in scan 4 and supporting the *Warning* mode. DS does not reflect at all the new available data, which are informative and should be taken into account. This pathological behavior could lead to wrong decisions. In our particular case however, DS leads to a right final decision at scan 10 by coincidence, but this decision could not be accepted as coherent and reliable, because it is not built on a consistent logical ground. Taking important decisions by chance could be critically wrong and could cause valuable damages.

For source 2, associated with *Alarm* mode (Fig. 4, middle subplot), Pign.Proba established by DS are as follows: $BetP(E) = 1$, $BetP(A) = 0$, $BetP(W) = 0$. During the scans 1 and 10, because of the conflicts in obtained

measurements, DS generates a totally wrong Pign.Proba $BetP(E) = 1.0$ assigned to *Emergency*, producing a hard decision for *Emergency* case. DS leads here to false alarm. That way video camera will be steered in wrong direction, which in reality is not the direction with highest priority. It means, that the true most dangerous direction for reaction will be ignored.

For source 3, associated with *Warning* mode (Fig. 4, subplot in the bottom), Pign.Proba established by DS are as follows: $BetP(E) = 1$, $BetP(A) = 0$ and $BetP(W) = 0$. Here the same false alarm situation is established as in source 2. Actually at scan 10 DS establishes totally wrong decisions for source 2 and source 3. The only right decision taken for source 1 is obtained by coincidence (because of not responding behaviour of the rule) and has no logical ground.

Decision taken by Dempster's rule at scan 20*:*

At scan 20, according to source 1, DS keeps its nonresponding behaviour, leading to right, but taken by coincidence decision. According to sensor 3 DS keeps the false alarm, as at scan 10. It succeeds to take a right decision for source 2, associated with *Alarm* mode, after a longer delay in reestablishing the belief masses for *Alarm*, in comparison with PCR5 rule.

Decision taken by Dempster's rule at scan 30*:*

At this scan DS succeeds to keep the right decision for source 2. However, it keeps performing as at scan 20, producing right, but logically ungrounded decision for source 1, and false alarm for source 3. Taking important decisions, concerning security, by chance, could be critically wrong and dangerous. Steering camera toward wrong direction, on the base of false alarm, could become critical too, because that way the proper camera response will be mistaken.

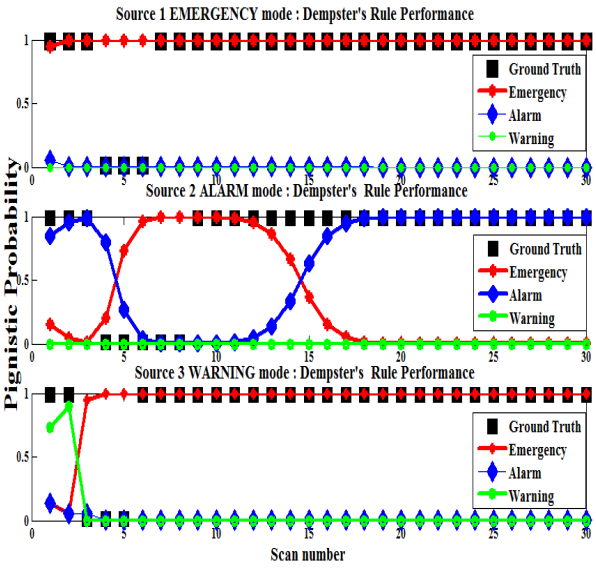


Fig. 4. Dempster's rule Performance for danger level estimation.

VII. Conclusions

In this paper the alarms' identification and prioritization (in terms of degree of danger) has been considered and realized using PCR5 rule of combination in order to estimate the proper degree of danger, especially in crowded scene, where events could happen at a set of a priori defined dangerous directions. The method utilized is based on the sequential fusion of the sound sources information obtained by two-dimensional microphone array defining the positions of the sources in surveillance area converted into basic belief assignments. A comparison of performance of PCR5 rule with respect to the performance of Dempster's rule has been done. The results obtained show the strong ability of PCR5 rule to take care in a coherent and stable way for the evolution of all possible degrees of danger, related to all the localized sources. It is especially significant in case of sound sources' data discrepancies and conflicts, when the highest priority mode *Emergency* occurs. PCR5 rule prevents to produce a mistaken decision, that way prevents to avoid the most dangerous case without immediate attention. A similar adequate behavior of performance is established in cases of lower danger priority. Dempster's rule shows weakness in resolving the cases examined. In *Emergency* case, Dempster's rule does not respond to the level of conflicts between sound sources, leading that way to ungrounded decisions. In cases of lower danger's priority (perturbed Warning and Alarm mode), Dempster's rule could cause a false alarm and can deflect the attention from the existing real dangerous source by assigning a wrong steering direction to the surveillance camera. In real world cases involving a broad surveillance area and multiple located sound sources, it becomes very important to realize distributed parallel processing with respect to the number of sources, in order to have correct decision in the proper time.

References

[1] Khosla R., Dillon T. Learning knowledge and strategy of a neuro expert system architecture in alarm processing, IEEE Trans. Power Systems, Vol. 12 (4), 1997, pp. 1610–1618.

[2] Vale Z., Machado A. An expert system with temporal reasoning for alarm processing in power system control centers, IEEE Trans. Power Systems, Vol. 8 (3), 1993, pp. 1307–1314.

[3] Lin W., Lin C., Sun Z. Adaptive multiple fault detection and alarm processing for loop system with probabilistic network, IEEE Trans. Power Delivery, Vol. 19 (1), pp. 64–69, 2004.

[4] Souza J., et al. . Alarm processing in electrical power systems through a neuro fuzzy approach, IEEE Trans. Power Delivery, Vol. 19 (2), pp. 537–544, 2004, .

[5] McArthur S., et al. The application of model based reasoning within a decision support system for protection engineers, IEEE Trans. Power Delivery, Vol. 11 (4), pp. 1748–1754, 1996.

[6] Foong O., Sulaiman S., Rambli D., Abdullah N. ALAP: Alarm Prioritization System For Oil Refinery, Proc. of the World Congr. on Eng. and Comp. Sci. Vol II, San Francisco, CA, USA, 2009.

[7] Smarandache F., Dezert J. (Editors) Advances and Applications of DSmT for Information Fusion, ARP, Rehoboth, Vol.1-3, 2004-2009.

[8] Behar V., et al. STAP Approach for DOA Estimation using Microphone Arrays, Signal Proc. Workshop, Vilnius, SPIE Proc., Vol. 7745, 2010.

[9] Shafer, G. A Mathematical Theory of Evidence, Princeton Univ., 1976.

[10] Smets P., Kennes R. The transferable belief model, Artif. Intel., 66 (2), pp. 191-234, 1994.

[11] Zadeh L. On the validity of Dempster's rule of combination, Memo M79/24, Univ. of California, Berkeley, USA, 1979.

Human Integration of Motion and Texture Information in Visual Slant Estimation

Albena Tchamova
Jean Dezert
Nadejda Bocheva
Pavlina Konstantinova

Originally published as Tchamova A., Dezert J., Bocheva N., Konstantinova P., *Human Integration of Motion and Texture Information in Visual Slant Estimation,* IEEE Intelligent Systems IS'2012, Sofia, Bulgaria, Sept. 6-8, 2012, and reprinted with permission.

***Abstract*—The present research is aimed to: *(i)* characterize the ability of human visual system to define the objects' slant on the base of combination of visual stimulus characteristics, that in general are uncertain and even conflicting. *(ii)* evaluate the influence of human age on visual cues assessment and processing; *(iii)* estimate the process of human visual cue integration based on the well known Normalized Conjunctive Consensus and Averaging fusion rules, as well on the base of more efficient probabilistic Proportional Conflict Redistribution rule no.5 defined within Dezert-Smarandache Theory for plausible and paradoxical reasoning. The impact of research is focused on the ability of these fusion rules to predict in adequate way the behavior of individuals, as well as age-contingent groups of individuals in visual cue integration process.**

***Keywords*—Integration of visual stimulus characteristics; DSmT; probabilistic Proportional Redistribution rule no.5; Normalized Conjunctive rule; Averaging rule.**

I. Introduction

The visual information about the 3D world utilized by humans is provided by a set of 2D images on the eye retina. It leads to uncertainty and/or discrepancy in image interpretations because the same projections could belong to different 3D objects. As an additional complexity, the visual system has to recover the information about objects' depth (i.e. the mutual disposition of objects) with respect to the observer. To overcome these difficulties one needs to utilize and combine in an effective way a variety of visual characteristics (or so called cues) in order to achieve inferences, more informative and potentially more accurate than if they were obtained by means of a single cue. The process of combining, manipulating and interpreting information in stimulus integration problem is beneficial because it allows the human visual system to estimate and perceive more accurately the objects' properties and to take appropriate actions, leading to improved reasoning (judgment) under uncertainty or/and possible conflicts between different visual stimulus. The uncertainty, associated with the utilized visual cues and the possible conflicts between them influences the decision making and action control in the process of human aging due to the increased level of internal noise in the visual system [1]. If the visual system neglects some of the available information [2], the visual signal/noise ratio will additionally deteriorate. Throughout the life cycle many aspects of vision and visual information processing decline and affect everyday task performance. 3D shape of objects and their spatial layout are specified on the base of both: static and dynamic visual cues. Age-related impairments in visual processing and perception are observed for both of them [3]. Therefore the task of vision inherently requires the integration of all available visual cue information to determine 3D object's shape. This paper focuses on human way of integration of motion and texture information in the process of object's slant estimation. Our goal is to reveal not only the age-related changes in the process of visual information assessment, but also the plasticity of the visual system to best adapt to these changes and to efficiently exploit all the available information in the visual scene in order to provide the visual system with a meaningful output, concerning more accurate and robust spatial information about the 3D objects. We will present and compare the performance of three fusion rules to model human way of visual cue integration: Normalized Conjunctive Consensus (NCC), Averaging (AVE), and the probabilistic Proportional Conflict Redistribution rule no.5 (pPCR5) defined recently within Dezert-Smarandache Theory (DSmT) for plausible and paradoxical reasoning. In section II we present briefly the visual cue integration problem and recall the principles of NCC, AVE and pPCR5 fusion rules. In section III, we present the experimental strategy and procedure, methods, subjects, involved in the experiments, stimulus and used apparatus. In section IV, the research reasoning logic is presented as well the results, obtained on the base of applied fusion rules. Concluding remarks are given in section V.

This paper is partially supported by contract $D002-240/18.12.2008$ with Bulgarian National Science Fund.

II. Visual Cue Integration for Slant Estimation

Vision provides a number of static and dynamic cues to the 3D layout of observed objects and scenes. Human show

individual differences in their abilities to utilize these cues for judgments. The first source of visual information considered in this paper relates to 2D texture variations in the projection of a slanted plane. The texture elements alter their form and the degree of shortening depends on their relative position in the plane and relative orientation to the observer - the shortening of element form and the texture density are highest in the direction of plane's tilt. The degree of texture variation depends on surface slant and it is biggest in the most distant plane areas with respect to the observer. Another source of visual information considers the object's motion relative to the observer. The gradient of velocity in two orthogonal directions contains information about the object's slant and tilt. When both static and motion information is available, the efficient way of combining data, provided by them, leads to more accurate and robust estimation of object's geometry and to better understanding and recognition of the surrounding scenes and objects. The common ideas for visual cue combination in order to specify the viewer-dependent object's characteristics rely on the assumption of cue independence. There are various methods for modeling the visual cue integration process. Bayesian inference [4], [5] is a classical approach for modeling and processing probabilistic information. An ideal Bayesian observer was used to define the optimal weighting and combination of redundant visual cues [8], [9]. The main difficulties applying it concern the need of measurements' statistics and knowledge about the a priori information. The Bayesian framework was applied for modeling the spatial integration of auditory and visual information [6], for visual and haptic integration [7] where the main idea is that the human brain combines visual cues to obtain the most reliable estimate of the state of the world, i.e. the estimate in which the variance of the resulting combined cue is minimized. As it will be shown in our research, this kind of integration, being very sensitive to the sources with the bigger means, neglects part of available information, which is very unsatisfactory behavior in cases of combining conflicting visual cues. Generally visual data are not only inaccurate, incomplete and uncertain, but even conflicting, because the observer moves, or the surfaces could change their orientation in the particular scene, or one object occludes the other. All these data particularities must be incorporated in the process of human visual perception in order to provide a complete and accurate model of the real world and to improve the decision accuracy. In our study we will apply and compare the performance of three fusion rules: NCC rule, pPCR5, and AVE fusion rules to model the human process of visual cues integration.

A. Normalized Conjunctive Consensus rule

The Normalized Conjunctive Consensus (NCC) rule is used to combine simultaneously assumed independent visual cues. In case considered, the information obtained by the available visual cues is characterized by Gaussian likelihood functions with given means $\mu_i, i = 1, 2, ..$ and standard deviations $\sigma_i, i = 1, 2, ..$, defining the uncertainty encountered in data. In case of two independent visual cues with one-dimensional Gaussian distributions $p_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp -\frac{1}{2}(\frac{x-\mu_1}{\sigma_1})^2$ and $p_2(x) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp -\frac{1}{2}(\frac{x-\mu_2}{\sigma_2})^2$, the combined distribution based on NCC rule becomes:

$$p_{NCC}(x) = \frac{1}{\sigma_{NCC}\sqrt{2\pi}} \exp -\frac{1}{2}(\frac{x-\mu_{NCC}}{\sigma_{NCC}})^2 \qquad (1)$$

$$\text{where} \quad \sigma^2_{NCC} = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2} \quad \text{and} \quad \mu_{NCC} = \sigma^2_{NCC}(\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}).$$

It is characterized with a mean, biased toward the function with the bigger of the two means, similar to Bayesian estimator. It is optimal (minimizes the variance of the error estimation), when the original distributions have close mean values. When two visual cues are in conflict, however, (characterized with distant distributions), NCC rule leads to neglecting (not utilizing) part of the available information, because the source with the bigger mean is weighted more heavily. In this case it is reasonable to keep the original distributions in the fused probability density function until it is possible to make reliable decision. This has been done by pPCR5 fusion rule defined in DSmT .

B. Probabilistic Proportional Conflict Redistribution rule no.5

The general principle of all Proportional Conflict Redistribution rules [10], Vol.3 is to: *1)* calculate the conjunctive consensus between sources of evidence (different visual cues) *2)* calculate the total or partial conflicting masses; *3)* redistribute the conflicting mass (total or partial) proportionally on non-empty sets involved in the model according to all integrity constraints. The recently proposed non-Bayesian probabilistic Proportional Conflict Redistribution rule no.5 [11] is based on the discrete Proportional Conflict Redistribution rule no.5 (PCR5) [10], Vol.3, for combining discrete basic belief assignments. For completeness, we will discuss in brief the main idea behind the discrete PCR5. It comes from the necessity to deal with both uncertain and conflicting information, transferring partial or total conflicting masses proportionally only to non-empty sets involved in the particular conflict and proportionally to their individual masses. Basic belief assignment (bba) represents the knowledge, provided by particular source of information about its belief in the true state of the problem under consideration. Given a frame of hypotheses $\Theta = \{\theta_1, ..., \theta_n\}$, and the so called power set $2^\Theta = \{\emptyset, \theta_1, ..., \theta_n, \theta_1 \cup \theta_2, ..., \theta_1 \cup \theta_2 \cup ... \cup \theta_n\}$, on which the combination is defined, the general basic belief assignment is defined as a mapping $m_s(.) : 2^\Theta \to [0, 1]$, associated with the given source of information s, such that: $m_s(\emptyset) = 0$ and $\sum_{X\in 2^\Theta} m_s(X) = 1$. The quantity $m_s(X)$ represents the mass of belief exactly committed to X. Under Shafer's model assumption of the frame Θ (requiring all the hypotheses to be exclusive and exhaustive), the PCR5 combination rule for only two sources of information is defined as: $m_{PCR5}(\emptyset) = 0$

and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{Y \in 2^\Theta \setminus \{X\} \\ X \cap Y = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)} \right] \quad (2)$$

All sets involved in the formula are in canonical form. The quantity $m_{12}(X)$ corresponds to the conjunctive consensus, i.e: $m_{12}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2)$. All denominators are different from zero. If a denominator is zero, that fraction is discarded. No matter how big or small the conflicting mass is, PCR5 mathematically does a proper redistribution of the conflicting mass since PCR5 goes backwards on the tracks of the conjunctive rule and redistributes the partial conflicting masses only to the sets involved in the conflict and proportionally to their masses put in the conflict, considering the conjunctive normal form of the partial conflict. PCR5 is quasi-associative and preserves the neutral impact of the vacuous belief assignment. The probabilistic PCR5 is an extension of discrete PCR5 version to its continuous probabilistic counterpart. Basic belief assignment, involved in discrete PCR5 rule is extended to densities of probabilities of random variables. For two independent sources of information with given Gaussian distributions $p_1(x)$ and $p_2(x)$, the obtained combined result becomes [11]:

$$p_{pPCR5}(x) = p_1(x) \int \frac{p_1(x) p_2(y)}{p_1(x) + p_2(y)} dy + p_2(x) \int \frac{p_2(x) p_1(y)}{p_2(x) + p_1(y)} dy \quad (3)$$

The behavior of pPCR5 fusion rule in comparison to NCC rule (1) could be characterized by two cases below:

Case 1: both densities $p_1(x)$ and $p_2(x)$ are close (Fig.1-case 1). The combined density acts as an amplifier of the information by reducing the variance. Here pPCR5 acts as NCC fusion rule.

Case 2: the densities $p_1(x)$ and $p_2(x)$ are distant (Fig.1-case 2). Then the combined density keeps both original densities (not merging both densities into only one unimodal Gaussian density as NCC rule does), avoiding to neglect a part of the available information.

This new (from a theoretical point of view) property is very interesting and it presents advantages for practical applications as it will be shown in our particular research. Application of

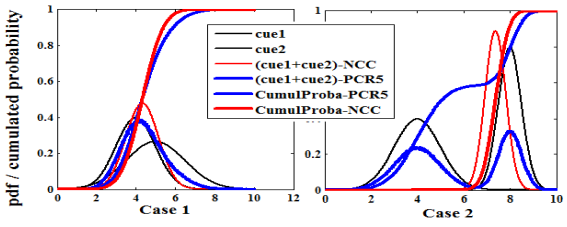


Fig. 1. Performance of pPCR5 fusion rule vs. NCC rule.

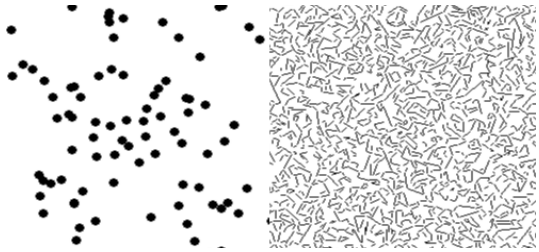


Fig. 2. Texture types: (left) dots, (right) lines.

pPCR5 fusion rule assures robustness to the potential errors and allows taking more reliable and adequate decisions in the process of integration of different cues in visual perception.

C. Averaging rule

The discrete simple Averaging rule consists in a simple arithmetic average of belief functions associated with sources of information (in our case particular visual cues). For given two sources of information defined with their discrete bba's: $m_1(.)$ and $m_2(.)$, for $\forall X \in 2^\Theta \setminus \{\emptyset\}$, the combined distribution based on this rule becomes $m_{AVE}(X) = \frac{1}{2}(m_1(X) + m_2(X))$. This trade-off rule is commutative, but not associative. In case of two independent and equally reliable/trustful visual characteristics, associated with Gaussian distributions: $p_1(x)$ and $p_2(x)$, the combined distribution based on Averaging rule becomes:

$$p_{AVE}(x) = \frac{1}{2}(p_1(x) + p_2(x)) \quad (4)$$

III. Experimental Goal, Methods, and Procedure

The experimental goal is directed to: (*i*) characterize the ability of human visual system to define the objects' slant on the base of only single cue available: Texture Information Only (referred as **TIO case**) or Motion Information Only (referred as **MIO case**), as well as in the case of both Texture and Motion information (referred as **TM case**), since human show significant individual differences in their abilities to combine and utilize both texture and motion information for judgments; (*ii*) evaluate the influence of human age on the assessment of objects' characteristics using available visual information.

A. Observers

Twelve elderly (mean age 74 years, range 67-85 years) and twelve younger (mean age 21 years, range 18-25 years) subjects took part in the experiments. All of them have passed eye examination. None of them reported having any major health problems.

B. Stimuli

The stimuli represent two slanted textured planes that form a symmetric horizontal dihedral angle. Two types of textures were rendered over the planes: dots (Fig.2-left) or a texture of non-intersecting lines (Fig.2-right).

Nine different sizes of the dihedral angles were used: 20 deg, 35 deg, 50 deg, 65 deg, 80 deg, 95 deg, 110 deg, 125 deg, and 140 deg. To change the size of the dihedral angle, the

slant of the two planes that hinged together was changed by an equal amount. One static and two dynamic conditions were generated. In all conditions the dihedral angle was presented in the middle of a computer screen under perspective projection and its vertical dimensions were fixed to 7 degrees of visual angle. In the static condition (**TIO** case) the information about the surface slant and consequently about the size of the dihedral angle is provided only by the changes in the texture over the planes. In the dynamic conditions (**MIO** case) the dihedral angle translated horizontally leftwards or rightwards with a speed of 6.4 deg of arc/s. It changed direction on every 1.1 s. In one of these conditions the texture specifies a flat object and thus, the information about the surface slant and the size of the dihedral angle is provided only by the motion. To achieve this, the texture coordinates were calculated relative to the eye coordinate system and they did not vary with the relative depth of the planes forming the dihedral angle. In the other dynamic condition (**TM** case) both the texture variation and the velocity of the object parts depend on the relative depth and therefore both specify the surface slant and the size of the dihedral angle.

C. Apparatus

The stimuli were presented on 21" Dell Trinitron monitor with Nvidia Quadro 900XGL graphic card. The monitor resolution was 1600 x 1200 pixels and the refresh rate was 85 Hz. The stimuli were rendered on the screen using OpenGL. Grayscale images with 8 bit precision (256 colors) were used. The monitor was gamma-corrected using a lookup table.

D. Experimental Procedure

The observer sat in semi-illuminated room at a distance of 114 cm from the computer screen. The method of single stimuli was used. On every trial the observers had to compare the stimulus with an internal standard - a right dihedral angle. The task of the observers was to evaluate whether the presented dihedral angle was larger/smaller than a right angle. Each subject participated in 6 sessions. The sessions differed by the experimental condition and the texture type. The order of the experimental sessions was contra-balanced across observers. In every experimental session the 9 different values of the dihedral angle were presented in random order 30 times. Each experimental session started with a demo to familiarize the subjects with the texture types (Fig. 2) used in the study and the way the texture changes in the different experimental conditions. The proportion of responses "the dihedral angle is larger than the right angle" is estimated for all different experimental conditions and for each subject the resulting psychometric functions are obtained. For example the observed psychometric function, associated with the first tested young subject for the case **TM** is given in Table I.

All subjects passed a priori training session of 60 trials in which a particular checkerboard pattern (Fig.3-left) was used to texture the dihedral angle under perspective projection (Fig.3-right). It helps the subjects to get familiar with the task to perform. The results of training were not taken into account.

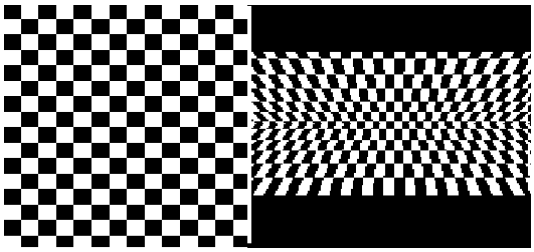


Fig. 3. Checkerboard pattern (left), Angle under perspective projection (right).

The dihedral angle remained visible on the screen until an answer was received. To give response the younger subjects used the buttons of a computer mouse while the elderly gave an oral response that was recorded by the experimenter.

IV. Experimental and Research Logic

Once having all psychometric functions, obtained for all different experimental conditions and for each subject in age-contingent groups, we should answer several questions:

Question 1. What is the effect of texture (*dots*, *lines*) in **MIO** case? Does the manipulation of the texture we applied succeed to eliminate it's effect in **MIO** case?

Question 2. How human observers combine the visual cues in order to estimate surface's slant? Do they base their responses on a single cue (**MIO** or **TIO**) or on their combination **TM**? If a single cue is used, which one - **TIO** or **MIO** is more informative?

Question 3. What combination rule (NCC, pPCR5, or AVE) used to combine available visual cues predicts more adequately human's way of cue integration?

Question 4. What is the common trend, concerning the *visual cue combination* performance of both age-contingent groups, i.e the performance of the so-called *averaged-people*, associated with each group. We denote these trends as: trend of *averaged-young-people* and respectively trend of *averaged-old-people*. They are based on combined individual behaviors in particular group, reveling its intrinsic behavior as a whole, reducing uncertainties associated with individual performances. All the tested subjects in age groups are considered as independent and equally reliable sources of information, because each subject provides his/her own psychometric function, associated with the **TM** combination process and should be taken into account in equal rights to derive these trends. Our goals are: *(i)* to find out which combinational rule (NCC, pPCR5 or AVE) is able to model correctly and adequately such human age-contingent group trends in reasoning process; *(ii)* to analyze the special features, characterizing these trends.

TABLE I
Example of psychometric function.

Angle's Value	20	35	50	65	80	95	110	125	140
Answers angle$> 90deg$ over 30 times	0	0.12	0.17	0.73	0.9	0.95	0.98	1	1

V. RESULTS

The experimental psychometric functions for both age groups and for all experimental conditions were compared using the *pfcmp* extension of MATLAB toolbox *psignfit* [14]. It implements a maximum-likelihood method [12] for fitting the psychometric functions and compares the parameters of the fits when estimated from the separate data sets and when the two sets are combined. As a result the significance value p is produces as a measure of fit between the examined psychometric functions.

• ***Results concerning Question 1 stated in Section IV.***

The results show that in **MIO** case the effect of the texture's type (*line* or *dots*) is effectively eliminated - for 10 out of 12 observers in each age group the null hypothesis of equal psychometric functions for both texture types could not be rejected at the assumed reliability level of $p = 0.05$. For static **TIO** case the comparison of the psychometric functions obtained for *line* texture and for *dots* texture for both age groups, shows that the null hypothesis could not be rejected at $p = 0.05$ for only 3 subject in each age group. These results suggest that the differences in the texture's type affect the subjects' performance significantly more in the static case. The smaller effect of the texture's types in **MIO** case provides indirect evidence that in these conditions the subjects' performance is based on the motion information.

• ***Results concerning Question 2 stated in Section IV.***

In order to answer this question, we have analyzed and compared the experimental psychometric functions obtained for each subject in both age-contingent groups given the following cases:

- {dots-based TIO vs. dots-based MIO vs. dots-based TM}
- {line-based TIO vs. line-based MIO vs. line-based TM}

Older people rely more on the static information, especially in case of *dots* texture type. Five out of 12 subjects do not show significant difference ($p = 0.05$) in their performance for the **TIO** and **TM** case for *dots* texture, and 4 out of 12 subjects - for *line* texture. Young people rely more on the dynamic information: the psychometric functions for **MIO** and **TM** case do not differ significantly at $p = 0.05$ for 5 out of 12 subjects for *dots* texture and 4 out of 12 - for *line* texture.

• ***Results concerning Question 3 stated in Section IV.***

In order to answer correctly this question we should evaluate the performances of applied combinational rules in the process of visual cue integration to predict the model of human fusion performance on the base of theoretically predicted psychometric functions. A comparison between experimentally obtained and predicted psychometric functions for all tested cases is provided on the base of *goodness-of-fit test* [13], one important application of chi-squared criteria: $\chi^2 = \sum_{j=1}^{J} \frac{(O_j - E_j)^2}{E_j}$ where χ^2 is an index of the agreement between an observed(O)/experimental and expected(E)/predicted via particular fusion rule sample values of psychometric function. For our case $J = 9$ represents the number of test angle values. The critical value of the test for $\nu = J - 1 = 8$ degrees of freedom at assumed $p = 0.01$ is $\chi^2 = 13.36$ [13]. This test is applied for both texture's types (*dots* and *line*) to the following pairs of psychometric functions:

- {MT case(experimental) - MT case (NCC rule)}
- {MT case(experimental) - MT case (pPCR5 rule)}
- {MT case(experimental) - MT case (AVE rule)}

TABLE II
CHI-SQUARED VALUES FOR OLDER SUBJECTS.

Subject	dotsTM pPCR5 / AVE	dotsTM NCC	lineTM pPCR5 / AVE	lineTM NCC
1	0.0653 / 0.0586	0.1359	0.1775 / 0.2032	0.5159
2	0.8415 / 0.9015	3.7232	0.0694 / 0.0663	0.0796
3	0.2359 / 0.2547	0.4360	0.0827 / 0.0934	0.1373
4	0.6995 / 0.6876	3.9117	0.1380 / 0.1461	0.1522
5	0.3618 / 0.3031	0.3751	0.2982 / 0.3098	0.3927
6	0.1066 / 0.1304	0.1387	0.1735 / 0.1943	0.2261
7	0.1859 / 0.1901	0.1935	0.1881 / 0.2101	0.4306
8	1.6944 / 1.7958	5.2330	0.3813 / 0.3114	0.4585
9	0.1697 / 0.2078	0.8814	1.0045 / 1.0062	1.5113
10	0.0368 / 0.0561	0.0566	0.1391 / 0.1411	0.1519
11	0.0909 / 0.0709	0.1021	0.0577 / 0.0499	0.0851
12	0.2664 / 0.2564	1.1320	0.1798 / 0.1682	1.5873

In general, the results show that the pPCR5 and AVE fusion rule predict more adequately than NCC rule human performance in all experimental cases. The differences between the experimental and estimated via pPCR5 and AVE rules psychometric functions for all observers in both age groups are smaller than those, obtained by NCC rule. For older subjects (Table II) all fusion rules predict psychometric functions that do not differ significantly from the experimental ones, but the differences in the fits are smaller in case of pPCR5 and AVE rules than in case of NCC rule application. For younger subjects (Table III), however, the NCC rule does not predict adequately the performance of the subjects in some conditions. For Subjects no. 5 and no. 6 (*dots-based* **TM** case) and for Subjects no. 4 and no. 9 (*lines-based* **TM** case) the obtained values (put in bold in Table III) significantly exceed the critical value of 13.36. The graphical results reflecting younger subjects' no. 4 and no. 9 fusion behaviors in *line* **TM** case are shown in Fig. 4. These results reflects the situations, when the experimentally obtained psychometric functions, associated with single cues (**TIO** and **MIO**) are characterized with distant underlying Gaussian distributions. In this case pPCR5 and AVE fusion rules make predictions, which model more correctly and adequately human fusion behavior. They are almost similar, but pPCR5 rule performs better than AVE rule in these conflicting cases. In the integration process, based on NCC rule however, part of available information was neglected, because the visual cues with bigger means were weighted more heavily (as it was described in Section II A.).

• ***Results concerning Question 4 stated in Section IV.***

In order to evaluate the common trend in the performance of both age groups, we started with the assumption that the tested subjects within each group are independent individual sources of information/answers and all of them are equally reliable. The results obtained for experimental and estimated (via different fusion rules) trends, concerning the *visual cue combination* groups' performance are presented in Fig. 5: subplots *1*, *3* for older group, and subplots *2*, *4* for younger one. Subplots *1* and *2* show results for *line* texture's type and subplots *3* and *4* - for *dots* texture's type.

In order to compare the performance of different fu-

TABLE III
CHI-SQUARED VALUES FOR YOUNGER SUBJECTS.

Subject	dotsTM pPCR5 / AVE	dotsTM NCC	lineTM pPCR5 / AVE	lineTM NCC
1	0.2976 / 0.3011	0.8526	0.0218 / 0.0191	0.0258
2	0.0801 / 0.0932	0.1456	0.1264 / 0.1525	0.6591
3	0.2182 / 0.2076	0.2690	0.1157 / 0.1201	0.1347
4	1.4509 / 1.4432	1.4716	0.6354 / 0.6523	**57.4916**
5	8.1655 / 8.1762	**45.1458**	1.4695 / 1.4512	2.4105
6	3.2425 / 3.3195	**34.1458**	0.1953 / 0.2003	12.2206
7	0.0014 / 0.0021	0.0079	0.2810 / 0.2957	0.9054
8	0.9201 / 0.8925	6.6588	0.3542 / 0.3513	0.9365
9	0.4950 / 0.4861	0.5160	0.8665 / 0.9341	**87.1105**
10	0.7633 / 0.7527	0.8304	0.1554 / 0.1599	0.1927
11	0.4202 / 0.4259	0.4380	0.3949 / 0.3901	0.3977
12	0.6371 / 0.6458	4.4540	0.0532 / 0.0525	0.2447

TABLE IV
CITY BLOCK ERRORS BETWEEN EXPERIMENTAL AND PREDICTED TRENDS.

	PCR5	NCC	AVErage
lineTM Older group	0.03	0.10	0.04
dotsTM Older group	0.06	0.11	0.04
lineTM Younger group	0.02	0.12	0.04
dotsTM Younger group	0.02	0.11	0.03

sion rules in estimating common trends' prediction the city-block errors between the corresponding pairs *averaged-young/old-people* MT(experimental) - *averaged-young/old-people* (NCC/pPCR5/AVE rules) for both texture's types are given in Table IV. Results show ultimatively that experimentally obtained trends and those, based on pPCR5 and AVE fusion rules are very closed and for both age-contingent groups are two times less then those, obtained via NCC fusion rule. pPCR5 and AVE rules predict more correctly the human model of reasoning, than NCC rule. pPCR5 performs a little bit better than AVE rule, utilizing all the available information (**TIO** and **MIO**), even in case of conflict. NCC based trends are very sensitive to the sources (different subjects' psychometric functions) with the bigger means, neglecting that way part of the available information and acting as an amplifier of the information by reducing the variances.

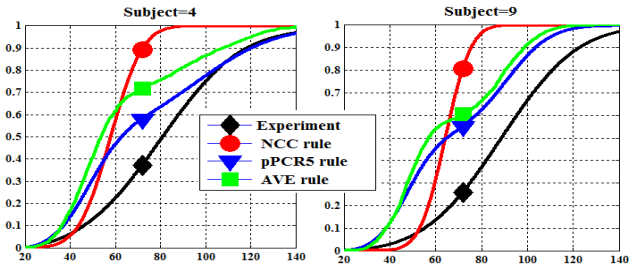


Fig. 4. Experimental and predicted performance for subject no.4 and no.9.

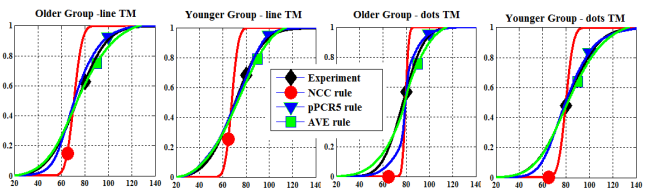


Fig. 5. Experimental and Predicted Trends in Performance of Age-related Groups.

VI. CONCLUSIONS

The results obtained in this study show age-related difference in the performance of the subjects in estimating the three-dimensional shape of the objects based on the texture and motion information. The task of the observers used in the study required the estimation of surface slant - a viewpoint dependent characteristic of the visual stimulation that is important for visual navigation and for object manipulation. Our data suggest that the younger people are more sensitive to differences in surface slant, but in the same time they are less accurate in their estimates. This cannot assure the robustness according to the potential errors during the experiments and leads to decisions which are less reliable than those taken by older people. Younger people as a group rely mainly on motion information neglecting the texture one. Elder people are characterized with less sensitivity to difference in the spatial characteristics of the three-dimensional objects in the real world, but they used to compensate this drawback by higher accuracy in their answers. Naturally this leads to ability to utilize correctly most of available stimulus information and then to improve the decision accuracy. The performance of both age groups in combining static and dynamic information is better described by the pPCR5 and AVE rule. In comparison to NCC rule, especially in conflicting cases pPCR5 fusion rules utilizes not only all available stimulus information, but this is achieved irrespective of the texture type (line or dots). That way pPCR5 fusion rule assures preserving the richness of stimulus data in the process of visual stimulus combination.

REFERENCES

[1] Pardhan, S. Contrast sensitivity loss with aging: sampling efficiency and equivalent noise at different spatial frequencies, JOSA A, Vol.21 (2), 2004, pp. 169–175.
[2] Falkenberg HK., Bex PJ. Sources of motion-sensitivity loss in glaucoma, Invest Ophthalmol Vis Sci., 48(6), 2007, pp. 2913–21.
[3] Owsley, C. Aging and vision, Vision Research, 2010.
[4] Bayes, T. An Essay towards solving a Problem in the Doctrine of Chances, Philosophical Transactions of the Royal Society of London, 1763, pp. 330–418.
[5] Sivia, D. Data Analysis, a Bayesian Tutorial, Clarendon (Oxford), 1996.
[6] Alais D., Burr D. The ventriloquist effect results from near optimal cross modal integration, Current Biology, 14, 2004, pp. 257–262.
[7] Ernst M., Banks M. Humans integrate visual and haptic information in a statistically optimal fashion, Nature, 415, 2002, pp. 429-433.
[8] Knill, D. Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant, Journal of Vision 7(7):5, 2007, pp. 1-24.
[9] Stocker A., Simoncelli E. Constraining a Bayesian Model of Human Visual Speed Perception, Advances in Neural Information Processing Systems, vol.17, 2005, pp. 1361–1368.
[10] Smarandache F., Dezert J. Advances and applications of DSmT for information fusion, Volumes 1, 2, 3, ARP, 2004–2009.
[11] Kirchner A., Dambreville F., Celeste F., Dezert J., Smarandache F. Application of probabilistic PCR5 Fusion Rule for Multisensor Target, In Proc.of International Conference of Information Fusion, Qubec, Canada, July 9-12, 2007.
[12] Wichmann FA., Hill NJ. The psychometric function: I. Fitting, sampling and goodness-of-fit. Perception and Psychophysics 63(8), 2001, pp. 1293–1313. Perception and Psychophysics 63(8), 2001, 1314-1329.
[13] Matre, J., Gilbreath G. Statistics for Business and Economics, 3rd Edition, ISBN 0-256-03719-1, 1987.
[14] Psignifit software, http://bootstrap-software.org/psignifit/.

Overview of Dempster-Shafer and Belief Function Tracking Methods

Erik Blasch
Jean Dezert
B Pannetier

Originally published as Blasch E., Dezert J., Pannetier *B., Overview of Dempster-Shafer and Belief Function Tracking Methods*, in Proc. SPIE, Vol. 8745, Baltimore, MD, USA, April 2013, and reprinted with permission.

ABSTRACT

Over the years, there have been many proposed methods in set-based tracking. One example of set-based methods is the use of Dempster-Shafer (DS) techniques to support belief-function (BF) tracking. In this paper, we overview the issues and concepts that motivated DS methods for simultaneous tracking and classification/identification. DS methods have some attributes, if applied correctly; but there are some pitfalls that need to be carefully avoided such as the redistribution of the mass associated with conflicting measurements. Such comparisons and applications are found in Dezert-Smarandache Theory (DSmT) methods from which the Proportional Conflict Redistribution (PCR5) rule supports a more comprehensive approach towards applying evidential and BF techniques to target tracking. In the paper, we overview two decades of research in the area of BF tracking and conclude with a comparative analysis of Bayesian, Dempster-Shafer, and the PCR5 methods.

Keywords: Dempster-Shafer, Belief Functions, DSmT, Target Recognition, Classification, & Identification, Tracking

1. INTRODUCTION

Humans and machines are typically trained for specific missions and/or scenarios [1]. One such case is classification of a moving target [2]. To integrate the benefits of human reasoning with machine methods, popular techniques of information fusion [3], target tracking [4], and pattern classification [5] are used.

When the human approaches the target, either the target is moving, the human is moving, or both are moving [6]. *Cognition*, the act of directing attention to sensory information, can be used by the human to fuse track and identify (ID) information as a perception of a set of moving targets. Dynamic cognitive multitarget-multisensor fusion under uncertainty requires target selection which can be formulated as a belief filtering problem in which sensed target states and identities are represented as current *situational beliefs*. The objective of the human is to 1) abstract number of tracks from the tracking environment, 2) assess confidence levels from the target classification algorithm, and 3) integrate the information for real-time beliefs of the number and types of targets from a plausible set of targets through an interactive display [7].

Multitarget tracking and ID is a subset of sensor fusion, which includes selecting sensors [8], sensor recognition policies, and tracking algorithms for a given set of mission requirements [9] for situational awareness [10, 11]. For example, in a typical tactical aircraft, the onboard sensors are active radar, electro-optical/infrared (EO/IR), and navigation sensors, with each sensor having a variety of modes in which it can operate and features it can detect. Figure 1 shows the case of EO/IR targets. The EO/IR sensor makes kinematic measurements to detect, track, and classify objects of interest while reducing user workload. In a dynamic and uncertain environment, a sensor manager, such as a human, must fuse the track and classification information to ID the correct target at a given time and can aid tracking algorithms by determining a set of tracks to follow and aid classification algorithms by constraining the set of plausible targets.

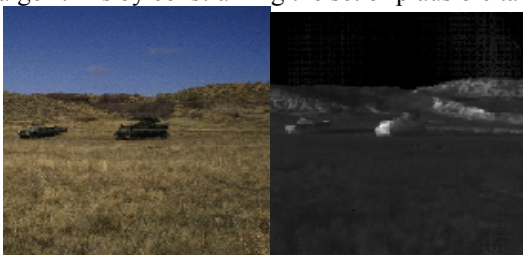


Figure 1. (a) Electro/Optical (EO) image and (B) Infrared (IR) image of targets [12].

Multitarget tracking in the presence of clutter has been investigated through the use of data association algorithms [1]. Likewise, other multisensor fusion algorithms have focused on tracking targets from multiple look sequences [13]. One inherent limitation of current algorithms is that the number of targets needs to be known *a priori*. While tracking algorithms can speculate on the number of targets, the cognition of the number of targets can be afforded to the user. The human is presented both the tracking information and the accrued evidence for each target type as a confidence measure. Once the human has an ID belief in the number of targets, the tracking algorithm can be updated. The human must then cognitively determine from the *set of targets* how many, what type, and which target goes with which track [14]. Additionally, the human has the ability to eliminate those targets that are not plausible which reduces the number of tracks and the set of pose templates from which the classification algorithm must search.

This paper presents a summary of DS methods in the last two decades. By introducing the operator or image analyst for cognitive fusion, it may offer a means to control some aspects of the computational burdens experienced by analytical data association techniques while improving track quality for multitarget tracking and ID in the presence of clutter. Section 2 overviews many applications of DS tracking. Section 3 describes DST methods while Section 4 compares the DST methods to Bayesian formulations. Section 5 presents a contemporary approach using the proportional conflict redistribution rule (PCR5) which is compared to the other methods in Section 6. Section 7 draws some conclusions.

2. TRACKING METHODS USING DEMPSTER-SHAFER THEORY

One of the earliest known works in applying Dempster-Shafer (DS) methods to target tracking was by Jean Dezert for navigation [15], where the sensor is moving and the targets are stationary. The emergence of the benefits of DS methods were applied by Robin Murphy for robotic scene analysis [16]. Building on Murphy's work, the DS methods were then applied to other robotic applications [17, 18, 19], albeit real-time control was still superior with Bayesian methods. At the same time, Johan Schubert applied DS methods for determining the number of tracks through the support and plausibility functions for the linking of submarine targets between tracks [20].

A few *tracking and identification algorithms* have been proposed for air-target tracking [21] and ground target tracking [22, 23] that extend Bayes' rule for identification where the most probable target is selected when there is incomplete knowledge. For instance, there are times when unknown targets might be of interest that are not known at algorithm initiation. At other times, there are unknown number of targets to track or targets not trained for classification. One way to study the problem is the interaction of the human and the machine working synergistically - since the sensors are extensions of the human's processing. The set theory approach to HRR target classification was proposed by Mitchell and Westerkamp and termed a *Statistical feature based classifier* (STaF) for air-target tracking [24]. In addition, Blasch [25, 26, 27] presented a feature-based set-theory approach for ground target tracking. In both cases, classification and tracking, a set of features and a set of targets was investigated by extending the STaF algorithm as a belief filter for radar profiles analysis from which a plausible set of tracks and targets are made available to the user at each time instant [28].

Given the ability to track individual targets using advances in DS theory for target identify and classification, methods were then developed for tracking a group of targets. In this case, the like targets could be grouped together based on common characteristics [29, 30, 31]. Additionally, Li [32] investigated methods for convex optimization for enhanced ID processing. Using a combination of belief filtering and data association improved analysis of maneuvering targets [33, 34]. Other group tracking methods were postulated for cluster-to-track fusion [35]. Finally, the benefits of DS methods provided complementary information to tracking through Kalman weighting [36], mutual aiding [37] and pose estimation [38].

Beginning in 2005, efforts were made to extend traditional DS tracking methods [39] to that of advanced techniques using the proportional conflict redistribution rule (PCR5) [40]. Other tracking methods included multisensor [41], activity analysis [42], and out-of-sequence methods [43]. Also, at that time, methods of combining DS with nonlinear tracking methods such as the particle filter [44] and the unscented Kalman filter [45] were developed. Finally, a fusion rule based on DS methods was used to solve the association problem in target tracking [46].

With the demonstrated performance of many applications of DS techniques, research continued in the exploration of DS rules for classification to improve track accuracy [47] and maneuvering targets [48]. Multisensor techniques were applied for heterogeneous sensor measurements [49], such as DS tracking with unattended ground sensor measurements [50]. Currently, efforts are sought for tracking performance evaluation with DS techniques for tracking and identification improvement [51, 52, 53]. Further assessment includes combinations with non-linear tracking methods [54] and

association of track segments [55]. More recent results have applied the developments from DS tracking from radar towards that of image processing [56, 57]. Throughout the many demonstrations of the successes of DS tracking, we now discuss DS methods, compare DST with Bayesian methods, and conclude with belief functions (BF) such as the contemporary PCR5 method for target tracking.

3. BASICS OF DEMPSTER-SHAFER (DST) THEORY

The Dempster-Shafer (DS) theory of evidence was devised as a means of dealing with *imprecise evidence* [58, 59] and has been applied to target classification [60]. Evidence concerning an unknown target is represented as a nonnegative set function $m : P(U) \rightarrow [0,1]$, where $P(U)$ denotes the set of subsets of the finite universe U such that $\mathrm{m}(\varnothing) = 0$ and $\Sigma_{S \subseteq U}\, \mathrm{m}(S) = 1$. The set function m is called a *mass assignment* and models a range of *possible beliefs* about propositional hypothesis of the general form $P_S \triangleq$ "object is in S_1" where $\mathrm{m}(S)$ is the *weight of belief* in the hypothesis P_S. The quantity $\mathrm{m}(S)$ usually interprets as the *degree of belief* that accrues in S, but to no proper subset of S. The weight of belief $\mathrm{m}(U)$ attached to the entire universe is called the *weight of uncertainty* and models our belief in the possibility that the evidence m in question is completely erroneous. The quantities

$$Bel_{\mathrm{m}}(S) \triangleq \sum_{O \subseteq S} \mathrm{m}(O) \tag{1}$$

$$Pl_{\mathrm{m}}(S) \triangleq \sum_{O \cap S \neq \varnothing} \mathrm{m}(O) \tag{2}$$

are called the ***belief*** and ***plausibility*** of the evidence, respectively and $\mathrm{m}(O)$ is the mass assignment for object O. The relationships $Bel_{\mathrm{m}}(S) \leq Pl_{\mathrm{m}}(S)$ and $Bel_{\mathrm{m}}(S) = 1 - Pl_{\mathrm{m}}(S^{\mathrm{c}})$ are true identically and the interval $[Bel_{\mathrm{m}}(S), Pl_{\mathrm{m}}(S)]$ is called the ***interval of uncertainty***, (IOU_{m}). Knowing that $Bel_{\mathrm{m}}(S) \rightarrow [0,1]$ and $Pl_{\mathrm{m}}(S) \rightarrow [0,1]$, three relationships exist. The first is a direct use of $Bel_{\mathrm{m}}(S)$ to accept and $Pl_{\mathrm{m}}(S)$ to reject measurements. Another method is to use the ***interval of certainty*** (IOC) defined as $[1, 1] - [Bel_{\mathrm{m}}(S), Pl_{\mathrm{m}}(S)]$ using interval subtraction. For example, [1, 1] - [0.8,0.9] = [1-0.9,1-0.8] = [0.1, 0.2]. Using the lower bound $Bel_{\mathrm{m}}(S)$ and upper bounds $Pl_{\mathrm{m}}(S)$ of the interval, we can assign a confidence measure $C_{\mathrm{m}} = 1 + Bel_{\mathrm{m}}(S) - Pl_{\mathrm{m}}(S) = 1 + 0.8 - 0.9 = 0.9$. Finally, the $IOU_{\mathrm{m}} = [Bel_{\mathrm{m}}(S), Pl_{\mathrm{m}}(S)]$ can be mapped to a Gaussian distribution for belief-based track tracking, using the IOU_{m} center as the mean, μ, rescaling the bounds to a Normal distribution and tacking the estimates (mean and variance). For example, $\mu_{\mathrm{m}} = (0.8+0.9)/2 = 0.85$. As another example, assume low belief with many measurements plausible, then $Bel_{\mathrm{m}}(S) = 0.3$ and $Pl_{\mathrm{m}}(S) = 0.9$, where $C_{\mathrm{m}} = 0.6$ is lower and the mean is $\mu_{\mathrm{m}} = 0.6$.

The mass assignment can be recovered from the belief function via the *Möbius transform*: [61]

$$\mathrm{m}(S) \triangleq \sum_{O \subseteq S} (-1)^{|S - O|}\, Bel_{\mathrm{m}}(O) \quad . \tag{3}$$

The set intersection quantity

$$(\mathrm{m} \oplus \mathrm{n})(S) \triangleq \frac{1}{1 - K} \sum_{X \cap Y = \mathrm{S}} \mathrm{m}(X(O))\, \mathrm{n}(Y(O)) \tag{4}$$

is called *Dempster's rule of combination*, where $K \triangleq \sum_{X(O) \cap Y(O) = \varnothing} \mathrm{m}(X(O))\, \mathrm{n}(Y(O))$ is called the ***conflict*** between the evidence m and evidence n.

In the finite-universe case, the Dempster-Shafer theory (DST) coincides with the theory of *independent, non empty* subsets of U (see [62, 63]; or for a dissenting view, see [64]). Given a mass assignment m, it is always possible to find a random subset Σ of U such that $\mathrm{m}(S) = p(\Sigma = S)$. In this case, $Bel_{\mathrm{m}}(S) = p(\Sigma \subseteq S) = \beta_{\Sigma}(S)$ and $Pl_{\mathrm{m}}(S) = p(\Sigma \cap S \neq 0) = \rho_{\Sigma}(S)$ where β_{Σ} and ρ_{Σ} are the belief and plausibility measures of Σ, respectively. Likewise, we can construct independent random subsets, Σ, Λ of U such that $\mathrm{m}(S) = p(\Sigma = \mathrm{S})$ and $\mathrm{n}(\mathrm{S}) = p(\Lambda = \mathrm{S})$ for all $S \subseteq U$. Then, it is easy to show that

$$(\mathrm{m} \oplus \mathrm{n})(S) = p(\Sigma \cap \Lambda \mid \Sigma \cap \Lambda \neq 0) \tag{5}$$

for all $S \subseteq U$. Thus, an intersection of overlapping sets can be fused to generate a global set confidence, where confidence is defined on the range [1, 1] - [Bel_m,Pl_m] and uncertainty is [Bel_m,Pl_m]. As comparative to the optimal approach, the next section provides a comparison of the set theory approach to that of traditional Bayesian analysis.

4. DEMPSTER-SHAFER VERSUS BAYESIAN THEORY

Recently, Dezert [65] has shown that Dempster's rule is consistent with probability calculus and Bayesian reasoning if and only if the prior $P(X)$ is uniform. However, when the $P(X)$ is not uniform, then Dempster's rule gives a different result. Both Yen [66] and Mahler [67, 68] developed methods to account for non-uniform priors. Others have also tried to compare Bayes and DST methods [69, 70, 71, 72, 73, 74, 75]. Assuming that we have multiple measurements $Z = \{Z_1, Z_2, \ldots\}$ for object O being tracked, Bayesian and DS methods are developed next.

Assuming conditional independence, one has the Bayes method:

$$\mathrm{P}(X \mid Z_1 \cap Z_2) = \frac{\mathrm{P}(X \mid Z_1)\,\mathrm{P}(X \mid Z_2) / \mathrm{P}(X)}{\sum_{i=1}^{N} \mathrm{P}(X_i \mid Z_1)\,\mathrm{P}(X_i \mid Z_2) / \mathrm{P}(X_i)} \tag{6}$$

With no information from Z_1 or Z_2, then $\mathrm{P}(X \mid Z_1, Z_2) = \mathrm{P}(X)$. Without Z_2, then $\mathrm{P}(X \mid Z_1, Z_2) = \mathrm{P}(X \mid Z_1)$ and without Z_1, then $\mathrm{P}(X \mid Z_1, Z_2) = \mathrm{P}(X \mid Z_2)$. Using Dezert's formulation, then the denominator can be expressed as a normalization coefficient:

$$m_{12}(\varnothing) = 1 - \sum_{X_i ; X_j \mid X_i \cap X_j} \mathrm{P}(X_i \mid Z_1)\,\mathrm{P}(X_i \mid Z_2) \tag{7}$$

Using this relation, then the total probability mass of the conflicting information is

$$\mathrm{P}(X \mid Z_1 \cap Z_2) = \frac{1}{1 - m_{12}(\varnothing)} \bullet \mathrm{P}(X \mid Z_1)\,\mathrm{P}(X \mid Z_2) \tag{8}$$

which corresponds to Dempster's rule of combination using Bayesian belief masses with uniform priors. When the prior's are not uniform, then Dempster's rule is not consistent with Bayes' Rule. For example, let $m_0(X) = P(X)$, $m_1(X) = P(X \mid Z_1)$, and $m_2(X) = P(X \mid Z_2)$, then

$$m(X) = \frac{m_0(X)\; m_1(X)\; m_2(X)}{1 - m_{012}(\varnothing)} = \frac{\mathrm{P}(X)\;\; \mathrm{P}(X \mid Z_1)\;\; \mathrm{P}(X \mid Z_2)}{\sum_{i=1}^{N} \mathrm{P}(X_i)\;\; \mathrm{P}(X_i \mid Z_1)\;\; \mathrm{P}(X_i \mid Z_2)} \tag{9}$$

Thus, methods are needed to deal with non-uniform priors and appropriately redistribute the conflicting masses.

5. DEZERT-SMARANDACHE THEORY (DSmT)

Recent advances in DS methods include *Dezert-Smarandache Theroy* (DSmT). DSmT is an extension to the Dempster-Shafer method of evidential reasoning which has been detailed in numerous papers and texts: *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3 [76]. In 2002, Dezert [77] introduced the methods for the reasoning and in 2003, presented the hyper power-set notation for DSmT [78]. Recent applications include the DSmT Proportional Conflict Redistribution rule 5 (PCR5) applied to target tracking. Other applications of DSmT can be found in the list of references at (http://www.onera.fr/staff/jean-dezert/) .

The key contributions of DSmT are the redistributions of masses such that no refinement of the frame Θ is possible unless a series of constraints are known. For example, Shafer's model [79] is the most constrained DSm hybrid model in DSmT.

Since Shafer's model, authors have continued to refine the method to more precisely address the combination of conflicting beliefs [80, 81, 82] and generalization of the combination rules [83, 84]. An adaptive combination rule [85] and rules for quantitative and qualitative combinations [86] have been proposed. Recent examples for sensor applications include electronic support measures, [87, 88], physiological monitoring sensors [89], and seismic-acoustic sensing [90].

Here we use the *Proportional Conflict Redistribution* rule no. 5 (PCR5) and no. 6 (PCR6) and the *Dezert-Smarandache Probability* (DSmP) selections which are discussed below. We replace Smets' rule [80] by the more effective PCR5 or eventually the more simple PCR6 and replace the pignistic transformation by the more effective DSmP transformation to estimate target classification probabilities. All details, justifications with examples on PCR5 and PCR6 fusion rules and DSmP transformation can be found freely from the web in the DSmT compiled texts [76], Vols. 2 & 3.

5.1. PCR5 and PCR6 fusion rules

In the DSmT framework, the PCR5 is used generally to combine the basic belief assignment (bba)'s. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. Let $m_1(.)$ and $m_2(.)$ be two independent bba's, then the PCR5 rule is defined as follows (see [76], Vol. 2 for full justification and examples): $m_{PCR5}(\varnothing) = 0$ and $\forall X \in 2^\Theta \setminus \{\varnothing\}$, where $\varnothing$ is the null set and 2^Θ is the power set:

$$m_{PCR5}(X) = \sum_{\substack{X_1;\, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) + m_2(X_2) \quad + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \varnothing}} \left[\frac{m_1(X_1)^2\, m_2(X_2)}{m_1(X_1) + m_2(X_2)} + \frac{m_1(X_1)\, m_2(X_2)^2}{m_1(X_1) + m_2(X_2)} \right] \qquad `(10)$$

where $\cap$ is the interesting and all denominators in the equation above are different from zero. If a denominator is zero, that fraction is discarded. Additional properties and extensions of PCR5 for combining qualitative bba's can be found in [76], Vol. 2 & 3. All propositions/sets are in a canonical form. A variant of PCR5, called PCR6 has been proposed by Martin and Osswald in [91], Vol. 2, for combining more than 2 sources. PCR6 coincides with PCR5 when one combines two sources. The difference between PCR5 and PCR6 lies in the way the proportional conflict redistribution is done as soon as three or more sources are involved in the fusion. For example, let's consider three sources with bba's $m_1(.)$, $m_2(.)$, and $m_3(.)$, $A \cap B = \varnothing$ for the model of the frame Θ, and $m_1(A) = 0.6$, $m_2(B) = 0.3$, and $m_3(B) = 0.1$. With ***PCR5*** the partial conflicting mass $m_1(A)\, m_2(B)\, m_3(B) = (0.6)(0.3)(0.1) = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR5} = 0.01714$ and $x_B^{PCR5} = 0.00086$ because the proportionalization is:

$$\frac{x_A^{PCR5}}{m_1(A)} = \frac{x_B^{PCR5}}{m_2(B)\, m_3(B)} = \frac{m_1(A)\, m_2(B)\, m_3(B)}{m_1(A) + m_2(B)\, m_3(B)} \qquad (11)$$

that is $$\frac{x_A^{PCR5}}{0.6} = \frac{x_B^{PCR5}}{(0.3)(0.1)} = \frac{0.018}{0.6 + 0.03} \approx 0.02857$$

thus $$x_A^{PCR5} = 0.60\ (0.02857) \approx 0.01714$$
$$x_B^{PCR5} = 0.03\ (0.02857) \approx 0.00086$$

With the ***PCR6*** fusion rule, the partial conflicting mass $m_1(A)\, m_2(B)\, m_3(B) = (0.6)(0.3)(0.1) = 0.018$ is redistributed back to A and B only with respect to the following proportions respectively: $x_A^{PCR6} = 0.0108$ and $x_B^{PCR6} = 0.0072$ because the PCR6 proportionalization is done as follows:

$$\frac{x_A^{PCR6}}{m_1(A)} = \frac{x_{B;2}^{PCR6}}{m_2(B)} = \frac{x_{B;3}^{PCR6}}{m_3(B)} = \frac{m_1(A)\, m_2(B)\, m_3(B)}{m_1(A) + m_2(B) + m_3(B)} \qquad (12)$$

that is

$$\frac{x_A^{PCR6}}{0.6} = \frac{x_{B;2}^{PCR6}}{0.3} = \frac{x_{B;3}^{PCR6}}{0.1} = \frac{0.018}{0.6 + 0.3 + 0.1} \approx 0.018$$

thus

$$x_A^{PCR6} = (0.6)\,(0.018) = 0.0108$$
$$x_{B,2}^{PCR6} = (0.3)\,(0.018) = 0.0054$$
$$x_{B,3}^{PCR6} = (0.1)\,(0.018) = 0.0018$$

and therefore with PCR6, one gets finally the following redistributions to A and B:

$$x_A^{PCR6} = (0.6)\,(0.018) = 0.0108$$

$$x_B^{PCR6} = x_{B,2}^{PCR6} + x_{B,3}^{PCR6} = 0.0054 + 0.0018 = 0.0072$$

From the implementation point of view, PCR6 is simpler to implement than PCR5. For convenience, Matlab codes of PCR5 and PCR6 fusion rules can be found in [76]. It is worth noting that there is a strong relationship between PCR6 rule and the averaging fusion rule which is commonly used to estimate the probabilities in the classical frequentist interpretation of probabilities. Such a probability estimate cannot be obtained using DS rule, nor the PCR5 rule and that is why we recommend to use PCR6 when combining more than two basic belief masses altogether [92].

5.2. The DSmP Transformation

The DSmP probabilistic transformation is an alternative to the classical pignistic transformation which allows us to increase the probabilistic information content (PIC), i.e. to minimize the Shannon entropy, of the approximated subjective probability measure drawn from any bba. Justification and comparisons of DSmP(.) with respect to BetP(.) and to other transformations can be found in details in [93, 76 Vol. 3, Chap. 3].

BetP: The pignistic transformation probability, denoted BetP, offers a compromise between maximum of credibility *Bel* and maximum of plausibility *Pl* for decision support. The *BetP* transformation is defined by $\mathrm{BetP}(\varnothing) = 0$ and $\forall X \in G^\Theta \setminus \{\varnothing\}$ by

$$BetP(X) = \sum_{Y \in G^\Theta} \frac{C_M(X \cap Y)}{C_M(Y)}\, m(Y) \tag{13}$$

where G^Θ corresponds to the hyper-power set including all the integrity constraints of the model (if any). $G^\Theta = 2^\Theta$ if one adopts Shafer's model for Θ and $G^\Theta = D^\Theta$ (Dedekind's lattice) if one adopts the free DSm model for Θ [76]. $C_M(Y)$ denotes the DSm cardinal of the set Y, which is the number of parts of Y in the Venn diagram of the model M of the frame Θ under consideration [76, Book 1, Chap. 7]. The BetP reduces to the *Transferable Belief Model* (TBM) when G^Θ reduces to classical power set 2^Θ when one adopts Shafer's model.

DSmP transformation is defined by $\mathrm{DSmP}_\varepsilon(\varnothing) = 0$ and $\forall X \in G^\Theta \setminus \{\varnothing\}$ by:

$$DSmP_\varepsilon(X) = \sum_{Y \in G^\Theta} \frac{\sum_{\substack{Z \subseteq X \cap Y \\ C(Z)=1}} m(Z) + \varepsilon \bullet C(X \cap Y)}{\sum_{\substack{Z \subseteq Y \\ C(Z)=1}} m(Z) + \varepsilon \bullet C(Y)}\, m(Y) \tag{14}$$

where $C(X \cap Y)$ and $C(Y)$ denote the cardinals of the sets $X \cap Y$ and Y respectively; $\varepsilon \geq 0$ is a small number which allows to reach a highest PIC value of the approximation of $m(.)$ into a subjective probability measure, and Z is the new evidence. Usually $\varepsilon = 0$, but in some particular degenerate cases, when the $\mathrm{DSmP}_{\varepsilon=0}(.)$ values cannot be derived, the $\mathrm{DSmP}_{\varepsilon>0}$ values can however always be derived by choosing ε as a very small positive number, say $\varepsilon = 1/1000$ for example in order to be as close as we want to the highest value of the PIC. The smaller ε, the better/bigger PIC value one gets. When $\varepsilon = 1$ and when the masses of all elements Z having $C(Z) = 1$ are zero, $\mathrm{DSmP}_{\varepsilon=1}(.) = \mathrm{BetP}(.)$.

6. COMPARATIVE RESULTS OF DS, BAYES, AND PCR5-BASED TRACKING

Here we simulate two scenarios of the three rules: Bayesian, Dempster-Shafer, and PCR5 rules of combination. For each scenario, we assume that the target information is collected from a sensor that is precise in the position measurements, but the uncertainty in either the sensor position accuracy or the classification information results in a confusion matrix (CM) formulation. With a two object representation being tracked (e..g, the standard fighter/cargo example), we have $CM = [O_1\ O_2; O_2\ O_1]$. In the first scenario, the target classification is $CM =$ `[0.75 0.25; 0.25 0.75]`, and the belief mass results are shown in Figure 2.

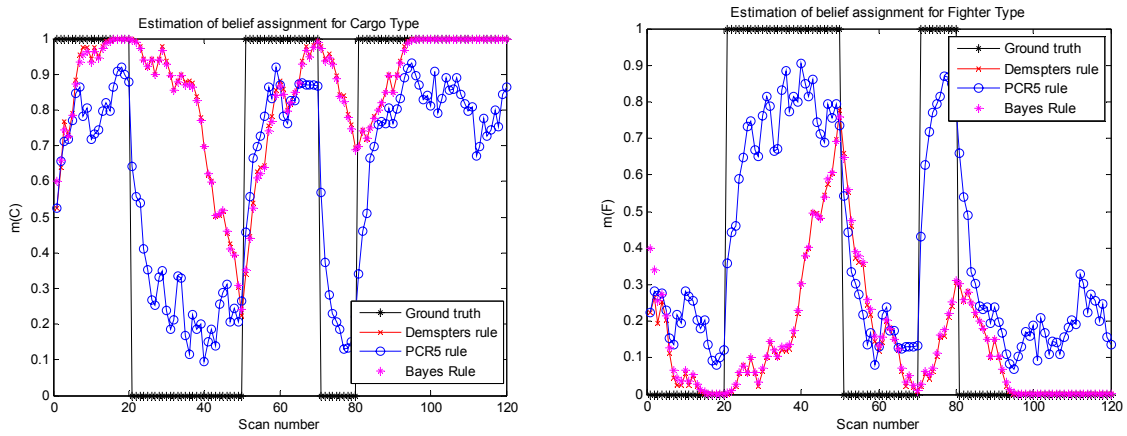


Figure 2. (a) Object 1 and (b) Object 2 tracking and identification results. $CM =$ `[0.75 0.25; 0.25 0.75]`

Figure 2 demonstrates that while there is uncertainty in the object tracking and classification, both the DS and Bayesian methods are close. The PCR5 results in better accuracy. Both DS and Bayesian methods have difficulty when the measurements change and suffer from *a prior* evidence biasing. In the next scenario, we decrease the sensor classification/ID accuracy; which results in more conflict in the analysis. For the Scenario 2 sensor model, we use $CM =$ `[0.65 0.35; 0.35 0.65]`.

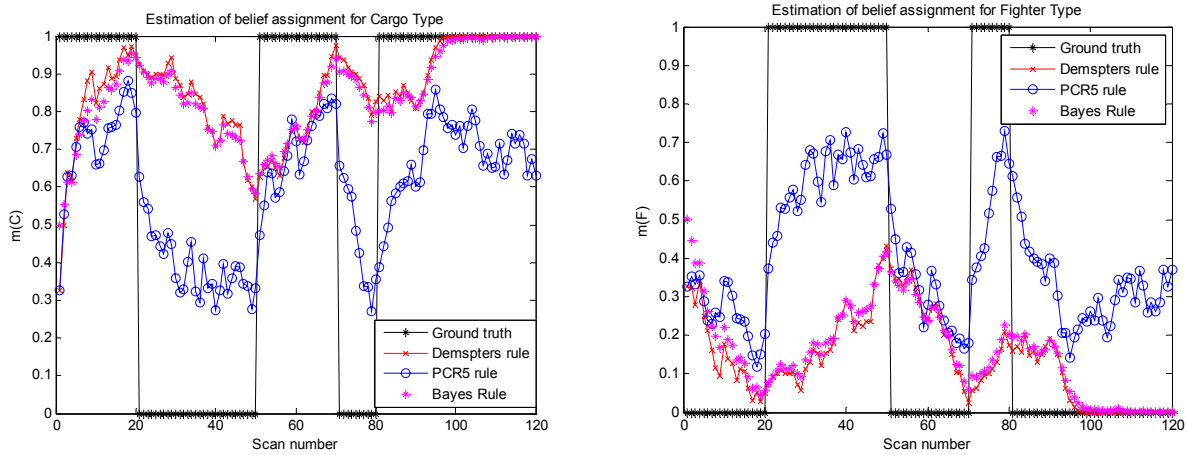


Figure 3. (a) Object 1 and (b) Object 2 tracking and identification results. $CM =$ `[0.65 0.35; 0.35 0.65]`.

Figure 3 illustrates differences between the three methods. DS tracking methods are able to improve over standard Bayesian methods when there is conflict in the measurements (Fig 3a scan 10 to 20). However, as shown in Figure 3, the PCR5 demonstrates an ability to track and ID the target when the measurement information is conflicting and changing (Fig 3a scan 25 to 50). The simple example illustrates the power of the PCR5 rule over standard DS and Bayesian methods to deal with conflicting, imprecise, and variations in target measurements for target tracking.

7. CONCLUSIONS

Conventional tracking techniques have difficulty in identifying targets when the number of targets is not known *a priori*., the targets are maneuvering, and there is conflict in the measurements. Throughout the last two decades, numerous researchers have explored Dempster-Shafer (DS) evidential (i.e., belief function) reasoning to solve the requirements of simultaneous tracking and identification. This paper has provided a literature review of most of the available publications that utilize the DS method in target, group/cluster, and multisensor tracking. Through a review of Bayesian, DS, and PCR5 formulations; we presented a simulated comparative example to demonstrate the current state-of-the-art methods such as DSmT research [94]. The PCR5 method can be extended to nonlinear tracking and ID algorithms, coordinated with users for assisted tracking, and can enhance conventional covariance and information filter formulations [95]. The presented PCR5 technique demonstrates promise for multitarget tracking problems and warrants further exploration with real-world data where environmental effects, occlusions, lost sensor data, and unknown targets [96] are standard.

REFERENCES

[1] Blasch, E. P., Bossé, E., and Lambert, D. A., [*High-Level Information Fusion Management and Systems Design*], Artech House, Norwood, MA, (2012).

[2] Kahler, B., and Blasch, E., "Decision-Level Fusion Performance Improvement from Enhanced HRR Radar Clutter Suppression," *J. of. Advances in Information Fusion,* Vol. 6, No. 2, Dec. (2011).

[3] Waltz, E., and Llinas, J., [*Multisensor Data Fusion*], Artech House, (1990).

[4] Bar-Shalom, Y., and Li, X. R., [*Multitarget-Multisensor Tracking: Principles and Techniques*], YBS, New York, (1995).

[5] Duda, R. O., Hart, P. E., and Stork, D. G., [*Pattern Classification* (2nd Edition)], Wiley, (2000).

[6] Blasch, E. P., and Watamaniuk, S. N., "Cognitive-based fusion using information sets for moving target recognition," *Proc. SPIE,* Vol. 4052, (2000).

[7] Blasch, E. P., "Assembling a distributed fused Information-based Human-Computer Cognitive Decision Making Tool," *IEEE Aerospace and Electronic Systems Magazine,* Vol. 15, No. 5, pp. 11-17, May (2000).

[8] Blasch, E., "Situation, Impact, and User Refinement," *Proc. of SPIE,* Vol. 5096, (2003).

[9] Blasch, E., "Level 5 (User Refinement) issues supporting Information Fusion Management" *Int. Conf. on Info Fusion*, (2006).

[10] Blasch, E., Kadar, I., Salerno, J., Kokar, M. M., Das, S., Powell, G. M., Corkill, D. D., and Ruspini, E. H., "Issues and Challenges in Situation Assessment (Level 2 Fusion)," *J. of Advances in Information Fusion,* Vol. 1, No. 2, pp. 122 - 139, Dec. (2006).

[11] Blasch, E., "Sensor, User, Mission (SUM) Resource Management and their interaction with Level 2/3 fusion" *Int. Conf. on Info Fusion*, (2006).

[12] Blasch, E., and Kahler, B., "Multi-resolution EO/IR Tracking and Identification" *Int. Conf. on Info Fusion*, (2005).

[13] Blasch, E., and Gainey, J., "Feature Based Biological Sensor Fusion," *Intl. Conference on Info. Fusion*, (1998).

[14] Blasch, E. P. and Plano, S. B., "JDL Level 5 Fusion model 'user refinement' issues and applications in group Tracking," *Proc. SPIE,* Vol. 4729, (2002).

[15] Dezert, J., "Autonomous navigation with uncertain reference points using the PDAF," in *Multitarget-Multisensor Tracking*, Vol 2, pp 271–324, Y. Bar-Shalom (Ed), Artech House, (1991).

[16] Murphy, R. R., "Dempster-Shafer Theory for Sensor Fusion in Autonomous Mobile Robots," *IEEE Tr. On Robotics and Automation*, Vol. 14, No. 2, April (1998).

[17] Blasch, E., Kondor, S., Gordon, M., and Hsu, R., "Georgia Tech Aerial Robotics team competition entry," *J. Aerial Unmanned Vehicle Systems*, pp. 20 – 25, May (1994).

[18] Blasch, E., "Flexible Vision-Based Navigation System for Unmanned Aerial Vehicles," *Proc. SPIE,* Vol. 2352, (1995).

[19] Balch, T., Boone, G., Collins, T., Forbes, H., MacKenzie, D., and Santamaria, J.C., "Io, Ganymede, and Callisto a multiagent robot trash-collecting team," *AI Magazine*, 16 (2): 39-53, (1995).

[20] Schubert, J, [*Cluster-based Specification Techniques in Dempster-Shafer Theory for an Evidential Intelligence Analysis of Multiple Target Tracks*], Ph.D. thesis, TRITA-NA-9410, Dept. of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, (1994).

[21] Leung, H., Li. Y., Bossé, E., Blanchette, M., Chan, K.C.C, "Improved multiple target tracking using Dempster-Shafer Identification," *Proc. SPIE*, Vol. 3068, (1997).

[22] Blasch, E., and Hong, L., "Simultaneous Tracking and Identification," *IEEE Conference on Decision Control*, pg. 249-256, (1998).

[23] Ding, Z., and Hong, L., "Decoupling probabilistic data association algorithm for multiplatform multisensor tracking," *Optical Engineering*, Vol. 37, No. 2, Feb. (1998).

[24] Micthell, R. A., and Westerkamp, J. J., "Robust Statistical Feature-based aircraft identification," *IEEE. TR. on Aerospace and Electronics Systems*, Vol. 35, No. 3, 1077-1094, (1999).

[25] Blasch, E., [*Derivation of a Belief Filter for Simultaneous High Range Resolution Radar Tracking and Identification*], Ph.D. Thesis, Wright State University, (1999).

[26] Blasch, E, Hong, L., "Sensor Fusion Cognition using belief filtering for tracking and identification," *Proc. SPIE,* Vol. 3719, (1999).

[27] Blasch, E. P., Westerkamp, J.J., Layne, J.R., Hong, L., Garber, F. D., and Shaw, A., "Identifying moving HRR signatures with an ATR Belief Filter," *Proc. SPIE,* Vol. 4053, (2000).
[28] Blasch, E., Hong, L., "Data Association through Fusion of Target track and Identification Sets," *Int. Conf. on Info Fusion*, (2000).
[29] Blasch, E., Connare, T., "Improving track accuracy through Group Information Feedback," *Int. Conf. on Info Fusion,* (2001).
[30] Blasch, E., Connare, T., "Improving Track maintenance Through Group Tracking," *Proc of the Workshop on Estimation, Tracking, and Fusion; A Tribute to Yaakov Bar Shalom*, Monterey, CA, 360 –371, May (2001).
[31] Connare, T., Blasch, E., Schmitz, J., Salvatore, F., Scarpino, F., "Group IMM tracking utilizing Track and Identification Fusion," *Proc. of the Workshop on Estimation, Tracking, and Fusion; A Tribute to Yaakov Bar Shalom,* 205 -220, May (2001).
[32] Li, J., Luo, Z-Q.,Wong, K.M., Bossé, E., "Convex optimization approach to identify fusion for multisensor target tracking," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, Volume: 31, pp. 172 – 178, (2001).
[33] Blasch, E., "Information-Theory-based feature-aided tracking and identification algorithm for tracking moving and stationary targets through high-turn maneuvers using fusion of SAR and HRR information", *Proc. of SPIE,* Vol. 4727, April (2002).
[34] Blasch, E., Connare, T., "Feature-Aided JBPDAF group tracking and classification using an IFFN sensor," *Proc. SPIE,* Vol. 4728, (2002).
[35] Schubert, J., "Robust report level cluster-to-track fusion," *Int'l Conf. on Information Fusion*, (2002).
[36] Wu, H., Siegel, M. ; Ablay, S., "Sensor fusion using Dempster-Shafer theory II: static weighting and Kalman filter-like dynamic weighting," *IEEE Proc. Instrumentation and Measurement Technology Conference*, (2003).
[37] Yang, C., and Blasch, E., "Mutual Aided Target Tracking and Identification," *Proc. of SPIE,* Vol. 5099, (2003).
[38] Yang, C., and Blasch, E., "Pose Angular-Aiding for Maneuvering Target Tracking", *Int. Conf. on Info Fusion*, (2005).
[39] Dezert, J., Tchamova, A. ; Semerdjiev, T., Konstantinova, P., "Performance evaluation of fusion rules for multitarget tracking in clutter based on generalized data association," *Int'l Conf. on Information Fusion*, (2005).
[40] Dezert, J., Tchamova, A., Smarandache, F., Konstantinova, P., "Target Type Tracking with PCR5 and Dempster's rules: A Comparative Analysis," *Int'l Conf. on Information Fusion*, (2006).
[41] Lancaster, J., Blackman, S, "Joint IMM/MHT Tracking and Identification for Multi-Sensor Ground Target Tracking," *Int'l Conf. on Information Fusion*, (2006).
[42] Snidaro, L., Piciarelli, C., Foresti, G.L., "Activity Analysis for Video Security Systems," *IEEE International Conference on Image Processing*, (2006).
[43] Maskell, S. R., Everitt, R. G., Wright, R., Briers, M., "Multi-target out-of-sequence data association: Tracking using graphical models," *Information Fusion*, Volume 7, Issue 4, Pages 434-447, December (2006).
[44] Faux, F., and Luthon, F., "Robust face tracking using color Dempster-Shafer fusion and particle filter," *Int'l Conf. on Information Fusion*, (2006).
[45] Khairnar, D.G., Nandakumar, S. , Merchant, S.N., Desai, U.B., "Nonlinear Target Identification and Tracking Using UKF," *IEEE Conf on Granular Computing*, (2007).
[46] Tchamova, A, Dezert, J., Smarandache, F., "A new class fusion rule for solving Blackman's Association Problem," *International IEEE Conference Intelligent Systems*, (2008).
[47] Kouemou, G., Neumann, C., Opitz, F., "Exploitation of track accuracy information in fusion technologies for radar target classification using Dempster-Shafer Rules," *Int'l Conf. on Information Fusion*, (2009).
[48] Dezert, J., Pannetier, B., "A PCR-BIMM filter for maneuvering target tracking," *Int. Conf. on Info Fusion*, (2010).
[49] Munz, M., Dietmayer, K., Mählisch, M., "Generalized fusion of heterogeneous sensor measurements for multi target tracking," *Int'l Conf. on Information Fusion*, (2010).
[50] Blasch, E., Maupin, P., Jousselme, A-L., "Sensor-Based Allocation for Path Planning and Area Coverage Using UGSs," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON)*, (2010).
[51] Blasch, E., Valin, P., "Track Purity and Current Assignment Ratio for Target Tracking and Identification Evaluation," *Int. Conf. on Info Fusion*, (2011).
[52] Pannetier, B., Dezert, J., "Extended and Multiple Target Tracking: Evaluation of an Hybridized Solution," *Int'l Conf. on Info Fusion*, (2011).
[53] Blasch, E., Straka, O., Yang, C., Qiu, D., Šimandl, M., Ajgl, J., "Distributed Tracking Fidelity-Metric Performance Analysis Using Confusion Matrices," *Int. Conf. on Info Fusion*, (2012).
[54] Liu, X., Leung, H., Valin, P., Bossé, E., "Multisensor joint tracking and identification using particle filter and Dempster-Shafer fusion," *Int'l Conf. on Information Fusion*, (2012).
[55] Pannetier, B., Dezert, J., "Track Segment Association with Classification Information," *Workshop on Sensor Data Fusion: Trends, solution, Applications*, 60-65, (2012).
[56] Dallil, A., Oussalah, M., Ouldali, A., "Sensor Fusion and Target Tracking Using Evidential Data Association," *IEEE Sensors Journal* Volume: 13 , Issue: 1, pp. 285 – 293, (2013).
[57] Li, X., Dick, A., Shen, C., Zhang, Z., Vandenhengel, A., Wang, H., "Visual Tracking with Spatio-Temporal Dempster-Shafer Information Fusion," *IEEE Tr. on Image Processing*, (2013).
[58] Buede, D., "Shafer-Dempster and Bayesian reasoning: A response to 'Shafer-Dempster reasoning with applications to multisensor target identification," *IEEE Transaction on Syst., Man and Cyber.,* vol.18, pp.1-10, (19880.
[59] Hall, D. L., [*Mathematical Techniques in Multisensor Data Fusion*], Artech House, (1992).
[60] Kruse, R., Schwencke, E., Heinsohn, J., [*Uncertainty and Vagueness in Knowledge-Based Systems*], Springer-Verlag, New York City, New York, (1991), update 2011.

[61] Mahler, R. P. S., "Random Sets in Information Fusion," in *Random Sets: Theory and Applications, Eds.* J. Goutsias, R.P.S. Mahler, H.T. Nguyen, IMA Volumes in Mathematics and its Applications, Vol. 97, Springer-Verlag Inc., New York, pp. 129-164, (1997).
[62] Hestir, K., Nguyen, H.T., Rogers, G.S., "A random set formalism for evidential reasoning," in *Conditional Logic in Expert Systems* (I.R. Goodman, N.M. Gupta, H.T. Nguyen, and G.S. Rogers, eds.) Amsterdam, The Netherlands: North-Holland, (1991).
[63] Nguyen, H. T., "On random sets and belief function," *J. of mathematical Analysis and Applications,* Vol 65, pp. 531-542, (1978).
[64] Smets, P., "The transferable belief model and random sets," *International J. of Intelligent Systems,* Vol 7, pp. 37 – 46, (1992).
[65] Dezert, J. "Non-Bayesian Reasoning for Information Fusion – A Tribute to Lofti Zadeh," *submitted to J. of Adv. of Information Fusion*, (2012).
[66] Yen, J., "A reasoning model based on the extended Dempster Shafer theory," *Nat Conf. on Artificial Intelligence*, (1986).
[67] Mahler, R.P., "Classification when a priori evidence is ambiguous," *Proc. SPIE*, Vol. 2234, (1994).
[68] Mahler, R.P., "Combining ambiguous evidence with respect to ambiguous a priori knowledge, I: Boolean logic," *IEEE Trans. Sys., Man & Cyber., Part A*, Vol. 26, pp. 27–41, (1996).
[69] Hong, L., Lynch, A, "Recursive temporal-spatial information fusion with applications to target identification," *IEEE Transactions on Aerospace and Electronic Systems*, Volume: 29 , Issue: 2, pp 435 – 445, (1993).
[70] Buede, D. M., Girardi, P., "A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, Volume: 27 , Issue: 5 , pp. 569 – 577, (1997).
[71] Blasch, E., "Learning attributes for situational awareness in the landing of an autonomous airplane," *AIAA/IEEE Digital Avionics Systems Conference*, (1997).
[72] Leung, H., Wu, J., "Bayesian and Dempster-Shafer target identification for radar surveillance," *IEEE Transactions on Aerospace and Electronic Systems*, Volume: 36 , Issue: 2, pp. 432 – 447, (2000).
[73] Schuck, T.M., Shoemaker, B., Willey, J., "Identification friend-or-foe (IFF) sensor uncertainties, ambiguities, deception and their application to the multi-source fusion process," *IEEE National Aerospace and Electronics Conference*, (2000).
[74] Mahler, R., "Can the Bayesian and Dempster-Shafer approaches be reconciled? Yes," *Int'l Conf. on Information Fusion*, (2005).
[75] Maskell, S. "A Bayesian approach to fusing uncertain, imprecise and conflicting information," *Information Fusion*, Vol. 9, 2, pp. 259-277, April (2008).
[76] Dezert, J. Smarandache, F., [*Advances and applications of DSmT for information fusion (Collected works)*], Vols. 1-3, American Research Press, 2004–2009. http://fs.gallup.unm.edu/DSmT.htm
[77] Dezert, J. "Foundations for a new theory of plausible and paradoxical reasoning," *Information & Security, An Int'l J.*, ed. by Prof. Tzv. Semerdjiev, Vol. 9, (2002).
[78] Dezert, J. Smarandache, F., "On the generation of hyper-powersets for the DSmT," *Int. Conf. on Info Fusion*, (2003).
[79] Shafer, G., [*A Mathematical Theory of Evidence*], Princeton, NJ: Princeton Univ. Press, (1976).
[80] Smets, P., "Analyzing the Combination of Conflicting Belief Functions," (2005).
[81] Daniel, M., "The MinC Combination of Belief Functions, derivation and formulas," *Tech Report No. 964*, Acad. Sci. of the Czech Republic, (2006).
[82] Jøsang, A., Daniel, M., "Strategies for Combining Conflict Dogmatic Beliefs," *Int. Conf. on Info Fusion*, (2006).
[83] Daniel, M., "Generalization of the Classic Combination Rules to DSm Hyper-Power Sets," *Information & Security, An Int'l J.*, Vol. 20, (2006).
[84] Smaradache, F., Dezert, J., "Information fusion based on new proportional conflict redistribution rules," *Int. Conf. Inf. Fusion*, (2005).
[85] Florea, M. C., Dezert, J., Valin, P., Smarandache, F., Jousselme, A-L., "Adaptive combination rule and proportional conflict redistribution rule for information fusion," *COGIS '06 Conf.*, (2006).
[86] Martin, A., Osswald, C., Dezert, J., Smarandache, F. "General Combination Rules for Qualitative and Quantitative Beliefs," *J. of Advances in Information Fusion*, Vol. 3, No. 2, Dec. (2008).
[87] Valin, P., Djiknavorian, P, Bossé, E., "A Pragmatic Approach for the use of Dempster-Shafer Theory in Fusing realistic Sensor Data," *J. of Advances in Info. Fusion*, Vol. 5, No. 1, June (2010).
[88] Djiknavorian, P., Grenier, D., Valin, P. "Approximation in DSm theory for fusing ESM reports," *Int. Workshop on Belief functions* 2010, Brest, France, April (2010).
[89] Lee, Z. H., Choir, J. S., Elmasri, R., "A Static Evidential Network for Context Reasoning in Home-Based Care," *IEEE Trans. Sys., Man, and Cyber-Part A; Sys & Humans*, Vol. 40, No. 6, Nov, (2010).
[90] Blasch, E., Dezert, J., Valin, P., "DSMT Applied to Seismic and Acoustic Sensor Fusion," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON)*, (2011).
[91] Martin, A., Osswald, C., "Generalized Proportion conflict redistribution rule applied to sonar imagery and radar targets classification," in *Adv. and Appl. of DSmT for Information Fusion*, Vol. 2, (Eds.) J. Dezert and F. Smarandache, (2006).
[92] Smarandache, F., Dezert, J., "On the consistency of PCR6 with the averaging rule and its application to probability estimation," *Int. Conf. on Info Fusion*, (2013).
[93] Dezert, J., Smarandache, F., "A new probabilistic transformation of belief mass assignment," *Int. Conf. on Info Fusion*, (2008).
[94] Yang, C., Kaplan, L., Blasch, E., "Performance Measures of Covariance and Information Matrices in Resource Management for Target State Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 48, No. 3, pp. 2594 – 2613, (2012).
[95] http://mmsip.bas.bg/mmosi/partners/jeandezert.html
[96] Mei, X., Ling, H., Wu, Y., Blasch, E., Bai, L. "Efficient Minimum Error Bounded Particle Resampling L1 Tracker with Occlusion Detection," *IEEE Trans. on Image Processing* (T-IP), (2013).

Tracking Applications with Fuzzy-Based Fusion Rules

Albena Tchamova
Jean Dezert

Originally published as Tchamova A., Dezert J., *Tracking applications with fuzzy-based fusion rules*, Proc. of 2013 IEEE International Symposium on INnovations in Intelligent SysTems and Application (INISTA 2013), Albena, Bulgaria, June 19-21, 2013, and reprinted with permission.

***Abstract*—The objective of this paper is to present and evaluate the performance of a particular fusion rule based on fuzzy T-Conorm/T-Norm operators for two tracking applications: (1) Tracking Object's Type Changes, supporting the process of identification, (e.g. friendly aircraft against hostile ones, fighter against cargo) and consequently for improving the quality of generalized data association; (2) Alarms identification and prioritization in terms of degree of danger relating to a set of a priori defined, out of the ordinary dangerous directions. The aim is to present and demonstrate the ability of TCN rule to assure coherent and stable way for identification and to improve decision-making process in temporal way. A comparison with performance of DSmT based PCR5 fusion rule and Dempster's rule is also provided.**

***Keywords*—Objects' type identification; Alarm classification; Data fusion; DSmT, TCN rule, PCR5 rule, Dempster's rule.**

I. Introduction

An important function of each surveillance system is to keep and improve targets tracks maintenance performance, as well as to provide a smart operational control, based on the intelligent analysis and interpretation of alarms coming from a variety of sensors installed in the observation area. Targets' type estimates can be used during different target tracking process stages for improving data to track association and for the quality evaluation of complicated situations characterized with closely spaced or/and crossing targets [1], [2]. It supports the process of identification, e.g. friendly aircraft against hostile ones, fighter against cargo. In such case, although the attribute of each target is invariant over time, at the attribute-tracking level the type of the target committed to the (unresolved) track varies with time and must be tracked properly in order to discriminate how many different targets are hidden in the same unresolved track. Alarms classification and prioritization [3],[4],[5],[6],[7],[8] is very challenging task, because in case of multiple suspicious signals (relating to a set of a priori defined, out of the ordinary dangerous directions), generated from a number of sensors in the observed area, it requires the most dangerous among them to be correctly recognized, in order to decide properly where the video camera should be oriented. There are cases, when some of the alarms generated could be incorrectly interpreted as false, increasing the chance to be ignored, in case when they are really significant and dangerous. That way the critical delay of the proper response could cause significant damages. In both cases above, the uncertainty and conflicts encountered in objects' and signals data, could weaken or even mistake the respective surveillance system decision. That is why a strategy for an intelligent, scan by scan, combination/updating of data generated is needed in order to provide the surveillance system with a meaningful output. In this paper we focus our attention on the ability of the so called T-Conorm-Norm (TCN) fusion rule, defined within Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning to improve the process of data fusion and to successfully finalize the decision-making procedures in both described surveillance cases. This work is based on preliminary research in [9],[10]. In section II we recall basics of Proportional Conflict Redistribution rule no.5 (PCR5), defined within DSmT. Basics of PCR5 based TCN fuzzy fusion rule are outlined in section III. Section IV presents the problem of alarms classification and examine the ability of TCN fusion rule to solve it. In section V the performance of TCN rule is analyzed related to the problem of target type tracking. In both sections, a comparative analysis of TCN rule solution with those, obtained by PCR5 and Dempster-Shafer's (DS) rule is provided. Concluding remarks are given in section VI.

II. Basics of PCR5 fusion rule

The general principle of Proportional Conflict Redistribution rules is to: $1)$ calculate the conjunctive consensus between the sources of evidences; $2)$ calculate the total or partial conflicting masses; $3)$ redistribute the conflicting mass (total or partial) proportionally on non-empty sets involved in the model according to all integrity constraints. The idea behind the Proportional Conflict Redistribution rule no. 5 defined within DSmT [9] (Vol. 2) is to transfer conflicting masses (total or partial) proportionally to non-empty sets involved in the model according to all integrity constraints. Under Shafer's model assumption of the frame Θ, PCR5 combination rule for only two sources of information is defined as: $m_{PCR5}(\emptyset) = 0$ and $\forall X \in 2^\Theta \setminus \{\emptyset\}$

$$m_{PCR5}(X) = m_{12}(X) + \sum_{\substack{X_2 \in 2^\Theta \setminus \{X\} \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(X_2)}{m_1(X) + m_2(X_2)} + \frac{m_2(X)^2 m_1(X_2)}{m_2(X) + m_1(X_2)}\right] \quad (1)$$

All sets involved in the formula (1) are in canonical form. $m_{12}(X)$ corresponds to the conjunctive consensus, i.e:

$$m_{12}(X) = \sum_{\substack{X_1, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2).$$

All denominators are different from zero. If a denominator is zero, that fraction is discarded. No matter how big or small is the conflicting mass, PCR5 mathematically does a better redistribution of the conflicting mass than Dempster-Shafer's rule since PCR5 goes backwards on the tracks of the conjunctive rule and redistributes the partial conflicting masses only to the sets involved in the conflict and proportionally to their masses put in the conflict, considering the conjunctive normal form of the partial conflict. PCR5 is quasi-associative and also preserves the neutral impact of the vacuous belief assignment.

III. Basics of TCN fusion rule

The T-Conorm-Norm rule of combination [11] represents a class of fusion rules based on specified fuzzy t-Conorm, t-Norm operators [16]. Triangular norms (t-norms) and Triangular conorms (t-conorms) are the most general families of binary functions that satisfy the requirements of the conjunction and disjunction operators, respectively. TCN rule is defined within DSmT based PCR5 fusion rule. Under Shafer's model assumption of the frame Θ, the TCN fusion rule for only two sources of information is defined as: $\tilde{m}_{TCN}(\emptyset) = 0$ and $\forall X \in 2^{\Theta} \setminus \{\emptyset\}$

$$\tilde{m}_{TCN}(X) = \tilde{m}_{12}(X) + \sum_{\substack{X_2 \in 2^{\Theta} \setminus \{X\} \\ X_2 \cap X = \emptyset}} \left[\frac{m_1(X).Tnorm\{m_1(X), m_2(X_2)\}}{Tconorm\{m_1(X), m_2(X_2)\}} + \frac{m_2(X).Tnorm\{m_2(X), m_1(X_2)\}}{Tconorm\{m_2(X), m_1(X_2)\}} \right] \quad (2)$$

where $\tilde{m}_{12}(X)$ corresponds to the conjunctive consensus, obtained by:

$$\tilde{m}_{12}(X) = \sum_{\substack{X_1, X_2 \in 2^{\Theta} \\ X_1 \cap X_2 = X}} Tnorm\{m_1(X_1), m_2(X_2)\}.$$

TCN fusion rule requires a normalization procedure :

$$\tilde{m}_{TCN}(X) = \frac{\tilde{m}_{TCN}(X)}{\sum_{\substack{X \in 2^{\Theta} \\ X \neq \oslash}} \tilde{m}_{TCN}(X)}$$

The attractive features of TCN rule could be defined as: very easy to implement, satisfying the impact of neutral Vacuous Belief Assignment; commutative, convergent to idempotence, reflects majority opinion, assures adequate data processing in case of partial and total conflict between the information granules. The general drawback of this rule is related to the lack of associativity, which is not a main issue in temporal data fusion.

IV. Alarms Classification Approach

The approach assumes all the localized sound sources to be subjects of attention and investigation for being indication of dangerous situations. The specific input sounds' attributes, emitted by each source, are sensor's level processed and evaluated in timely manner for their contribution towards correct alarms' classification (in term of degree of danger). The applied algorithm considers the following steps:

- Defining the frame of expected hypotheses as follows: $\Theta = \{\theta_1 = (E)mergency, \theta_2 = (A)larm, \theta_3 = (W)arning\}$. Here Shafer's model holds and we work on the power-set: $2^{\Theta} = \{\emptyset, E, A, W, E \cup A, E \cup W, A \cup W, E \cup A \cup W\}$. The hypothesis with a highest priority is *Emergency*, following by *Alarm* and then *Warning*.
- Defining an input rule base to map the sounds' attributes (so called *observations*) obtained from all localized sources into non-Bayesian basic belief assignments $m_{obs}(.)$.
- At the very first time moment $k = 0$ we start with a priori basic belief assignment (history) set to be a vacuous belief assignment $m_{hist}(E \cup A \cup W) = 1$, since there is no information about the first detected degree of danger according to sound sources.
- Combination of currently received measurement's bba $m_{obs}(.)$ (for each of located sound sources), based on the input interface mapping, with a history's bba, in order to obtain estimated bba relating to the current degree of danger $m(.) = [m_{hist} \oplus m_{obs}](.)$. TCN rule is applied in the process of temporal data fusion to update bba's associated with each sound emitter.
- Flag for an especially high degree of danger has to be taken, when during the a priori defined scanning period, the maximum Pignistic Probability [9] is associated with the hypothesis *Emergency*. In this work, we assume Shafer's model and we use the classical Pignistic Transformation [9], [15] to take a decision about the mode of danger. It is defined for $\forall A \in 2^{\Theta}$ by

$$BetP(A) = \sum_{X \in D^{\Theta}} \frac{|X \cap A|}{|X|} \cdot m(X) \quad (3)$$

where $|X|$ denotes the cardinality of X.

A. Simulation Scenario

A set of three sensors located at different distances from the microphone array are installed in an observed area for protection purposes, together with a video camera [13]. They are assembled with alarm devices: Sensor 1 with *Sonitron*, Sensor 2 with *E2S*, and Sensor 3 with *System Sensor*. In case of alarm events (smoke, flame, intrusion, etc.) they emit powerful sound signals with various duration and frequency of intermittence (Table 1), depending on the nature of the event.

Table 1 Sound signal parameters.

Continuous (*Warning*)	Intermittent-I (*Alarm*)	Intermittent-II (*Emergency*)
$f_{int} = 0$Hz	$f_{int} = 5$Hz	$f_{int} = 1$Hz
$T_{sig} = 10$s	$T_{sig} = 30$s	$T_{sig} = 60$s

The frequency of intermittencies f_{int}, associated with the localized sound sources is utilized in the specific input interface (the rule base) below.

Rule 1: if $f_{int} \rightarrow 1Hz$ then $m_{obs}(E) = 0.9$ and $m_{obs}(E \cup A) = 0.1$.

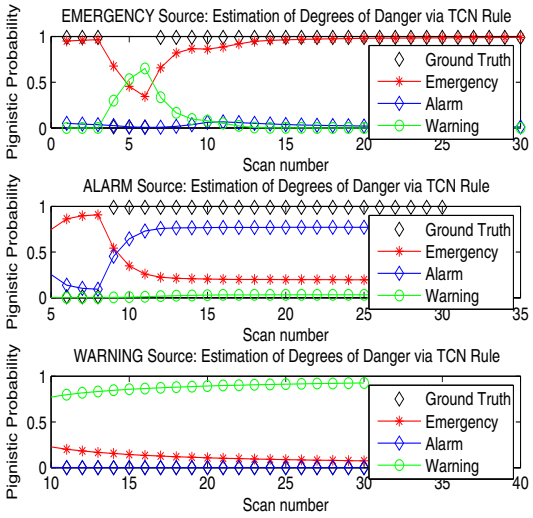

EMERGENCY Source: Estimation of Degrees of Danger via TCN Rule
ALARM Source: Estimation of Degrees of Danger via TCN Rule
WARNING Source: Estimation of Degrees of Danger via TCN Rule
Pignistic Probability
Scan number
Ground Truth
Emergency
Alarm
Warning

Fig. 1. TCN rule Performance for danger level estimation.

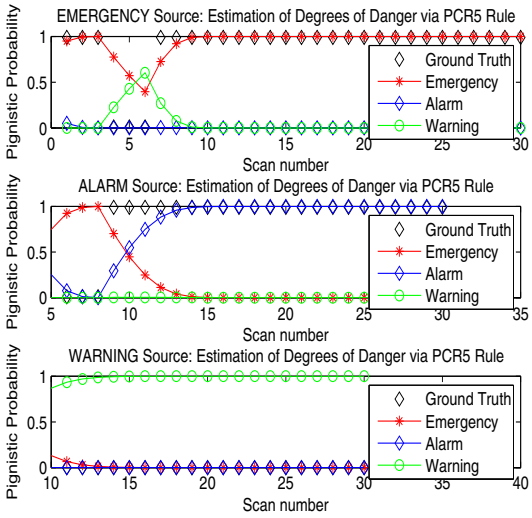

EMERGENCY Source: Estimation of Degrees of Danger via PCR5 Rule
ALARM Source: Estimation of Degrees of Danger via PCR5 Rule
WARNING Source: Estimation of Degrees of Danger via PCR5 Rule
Pignistic Probability
Scan number
Ground Truth
Emergency
Alarm
Warning

Fig. 2. PCR5 rule Performance for danger level estimation.

Rule 2: if $f_{int} \to 5Hz$ then $m_{obs}(A) = 0.7$, $m_{obs}(A \cup E) = 0.2$ and $m_{obs}(A \cup W) = 0.1$.

Rule 3: if $f_{int} \to 0Hz$ then $m_{obs}(W) = 0.6$ and $m_{obs}(W \cup A \cup E) = 0.4$.

Three main cases are estimated: the probabilities of modes, evaluated for Sensor 1 (associated with *Emergency* mode), Sensor 2 (associated with *Alarm* mode), and Sensor 3 (associated with *Warming* mode. The decisions should be governed at the video camera level, taken periodically, depending on: *1)* specificities of the video camera (time needed to steer the video camera toward a localized direction); *2)* time duration needed to analyze correctly and reliably the sequentially gathered information. We choose as a reasonable sampling period for camera decisions $T_{dec} = 20sec$, i.e. at every 10th scan.

B. TCN rule performance for danger level estimation.

Fig.1 shows the values of Pignistic Probabilities of each mode (*E*, *A*, *W*) associated with three sound emitters (1st source in *E* mode, (subplot on the top), 2nd source in *A* mode (subplot in the middle), and 3rd source in *W* mode, (subplot in the bottom)) during the all 30 scans. Each source has been perturbed with noises in accordance with the simulated *Ground Truth*, associated with particular sound source. These probabilities are obtained for each source independently as a result of sequential data fusion of $m_{obs}(.)$ sequence using TCN combinational rule. For a completeness of study and for comparison purposes, the respective performances of PCR5 and DS rule are presented in fig.2 and fig.3.

TCN rule shows a stable, quite proper and effective behavior, following the performance of PCR5 rule. A special feature of TCN rule performance are the smoothed estimates and more cautious decisions taken at the particular decisive scans.

The results obtained show the strong ability of PCR5 rule to take care in a coherent and stable way for the evolution of all possible degrees of danger, related to all the localized sources. It is especially significant in case of sound sources data

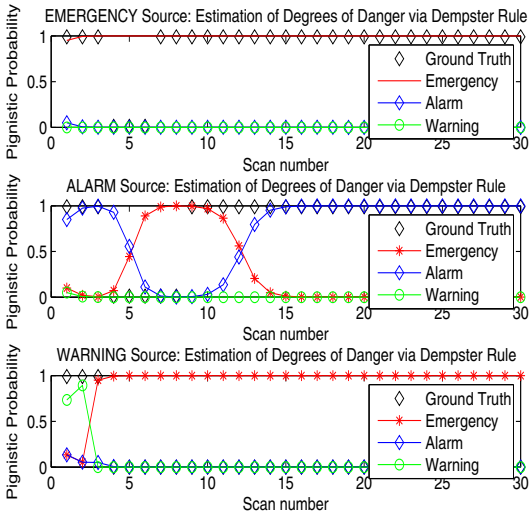

EMERGENCY Source: Estimation of Degrees of Danger via Dempster Rule
ALARM Source: Estimation of Degrees of Danger via Dempster Rule
WARNING Source: Estimation of Degrees of Danger via Dempster Rule
Pignistic Probability
Scan number
Ground Truth
Emergency
Alarm
Warning

Fig. 3. Dempster's rule Performance for danger level estimation.

discrepancies and conflicts, when the highest priority mode Emergency occurs. PCR5 rule prevents to produce a mistaken decision, that way prevents to avoid the most dangerous case without immediate attention. A similar adequate behavior of performance is established in cases of lower danger priority.

DS rule shows weakness in resolving the cases examined. In Emergency case, DS rule does not reflect at all new obtained informative observations supporting the Warning mode. This pathological behavior reflects the dictatorial power of DS rule realized by a given source [12], which is fundamental in Dempster-Shafer reasoning [14]. In our particular case however, DS rule leads to a right final decision by coincidence, but this decision could not be accepted as coherent and reliable, because it is not built on a consistent logical ground. In cases of lower dangers priority (perturbed Warning and Alarm mode), DS rule could cause a false alarm and can deflect the attention

from the existing real dangerous source by assigning a wrong steering direction to the surveillance camera.

V. Target Type Tracking Approach

The problem can be simply stated as follows:

- Let $k = 1, 2, ..., k_{max}$ be the time index and consider M possible target types $T_i \in \Theta = \{\theta_1, \ldots, \theta_M\}$ in the environment; for example $\Theta = \{Fighter, Cargo\}$ and $T_1 \triangleq Fighter$, $T_2 \triangleq Cargo$; or $\Theta = \{Friend, Foe, Neutral\}$, etc.
- at each instant k, a target of true type $T(k) \in \Theta$ (not necessarily the same target) is observed by an attribute-sensor (we assume a perfect target detection probability here).
- the attribute measurement of the sensor (say noisy Radar Cross Section for example) is then processed through a classifier which provides a decision $T_d(k)$ on the type of the observed target at each instant k.
- The sensor is in general not totally reliable and is characterized by a $M \times M$ confusion matrix

$$\mathbf{C} = [c_{ij} = P(T_d = T_j | TrueTargetType = T_i)]$$

The goal is to estimate $T(k)$ from the sequence of declarations done by the unreliable classifier up to time k, i.e. how to build an estimator $\hat{T}(k) = f(T_d(1), T_d(2), \ldots, T_d(k))$ of $T(k)$. The principle of the estimator is based on the sequential combination of the current basic belief assignment (drawn from classifier decision, i.e. our *measurements*) with the prior bba estimated up to current time from all past classifier declarations.

The algorithm follows the next main steps:

- Initialization step (i.e. $k = 0$). Select the target type frame $\Theta = \{\theta_1, \ldots, \theta_M\}$ and set the prior bba $m^-(.)$ as vacuous belief assignment, i.e $m^-(\theta_1 \cup \ldots \cup \theta_M) = 1$ since one has no information about the first target type that will be observed.
- Generation of the current bba $m_{obs}(.)$ from the current classifier declaration $T_d(k)$ based on attribute measurement. At this step, one takes $m_{obs}(T_d(k)) = c_{T_d(k)T_d(k)}$ and all the unassigned mass $1 - m_{obs}(T_d(k))$ is then committed to total ignorance $\theta_1 \cup \ldots \cup \theta_M$.
- Combination of current bba $m_{obs}(.)$ with prior bba $m^-(.)$ to get the estimation of the current bba $m(.)$. Symbolically we will write the generic fusion operator as $\oplus$, so that $m(.) = [m_{obs} \oplus m^-](.) = [m^- \oplus m_{obs}](.)$. The combination $\oplus$ is done according either Demspter's rule (i.e. $m(.) = m_D(.)$) or PCR5 rule (i.e. $m(.) = m_{PCR5}(.)$).
- Estimation of True Target Type is obtained from $m(.)$ by taking the singleton of Θ, i.e. a Target Type, having the maximum of belief (or eventually the maximum Pignistic Probability).
- set $m^-(.) = m(.)$; do $k = k + 1$ and go back to step b).

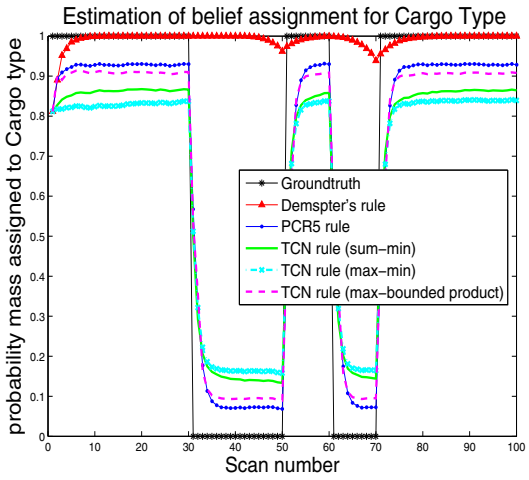


Fig. 4. Estimation of belief assignment for Cargo type.

A. Simulations results

In order to evaluate the performances of TCN-based estimator, a set of Monte-Carlo simulations on a very simple scenario for a 2D Target Type frame, i.e. $\Theta = \{(F)ighter, (C)argo\}$ is realized for classifier with a following confusion matrix:

$$\mathbf{C} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$$

We assume there are two closely spaced targets: Cargo and Fighter. Due to circumstances, attribute measurements received are predominately from one or another and both targets generates actually one single (unresolved kinematics) track. To simulate such scenario, a Ground Truth sequence over 100 scans was generated. The sequence starts with the observation of a Cargo type and then the observation of the target type switches two times onto Fighter type during different time duration. At each time step k the decision $T_d(k)$ is randomly generated according to the corresponding row of the confusion matrix of the classifier given the true target type (known in simulations). Then the algorithm from above is applied. The simulation consists of 10000 Monte-Carlo runs. The computed averaged performances (on the base of estimated belief masses obtained by the tracker) are shown on the figures 4 and 5. They are based on TCN fusion rule realized with different t-conorm and t-norm functions. On the same figures, for a comparison purposes, the respective performances of PCR5 and DS rule are presented. It is evident, that PCR5 fusion rule outperforms the results based on TCN rule, because PCR5 allows a very efficient Target Type Tracking, reducing drastically the latency delay for correct Target Type decision. TCN fusion rule shows a stable and adequate behavior, characterized with more smoothed process of re-estimating the belief masses in comparison to PCR5. TCN fusion rule with t-conorm=max and t-norm=bounded product reacts and adopts better than TCN with t-conorm=sum and t-norm=min, followed by TCN with t-conorm=max and t-norm=min.

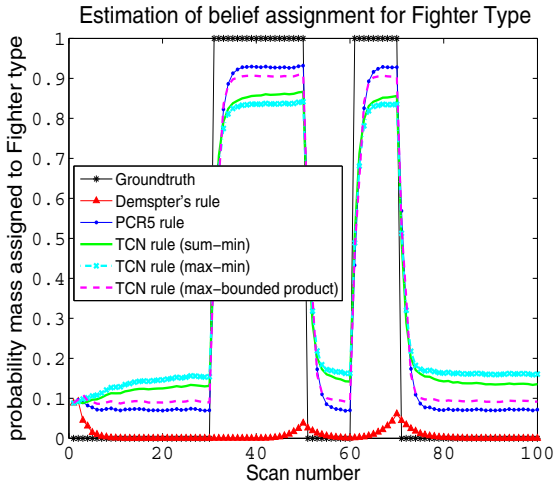


Fig. 5. Estimation of belief assignment for Fighter type.

presented: (1) Tracking Object's Type Changes, supporting the process of identification; (2) Alarms identification and prioritization in terms of degree of danger relating to a set of a priori defined, out of the ordinary dangerous directions. The ability of TCN rule to assure coherent and stable way of identification and to improve decision-making process in temporal way are demonstrated. Different types of t-conorm and t-norms, available in fuzzy set/logic theory provide us with richness of possible choices to be used applying TCN fusion rule. The attractive features of TCN rule is it's easy implementation and adequate data processing in case of conflicts between the information granules.

VII. Acknowledgment

The research work reported in the paper is partly supported by the project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions).

References

[1] Bar-Shalom Y., *Multitarget-Multisensor Tracking: Advanced Applications*, Artech House,1990.

[2] Blackman S., Popoli R., *Design and Analysis of Modern Tracking Systems*, Artech House,1999.

[3] Khosla R., Dillon T. Learning knowledge and strategy of a neuro expert system architecture in alarm processing, IEEE Trans. Power Systems, Vol. 12 (4), pp. 1610–1618, 1997.

[4] Vale Z., Machado A. An expert system with temporal reasoning for alarm processing in power system control centers, IEEE Trans. Power Systems, Vol. 8 (3), pp. 1307–1314, 1993.

[5] Lin W., Lin C., Sun Z. Adaptive multiple fault detection and alarm processing for loop system with probabilistic network, IEEE Trans. Power Delivery, Vol. 19 (1), pp. 64–69, 2004.

[6] Souza J., et al. . Alarm processing in electrical power systems through a neuro fuzzy approach, IEEE Trans. Power Delivery, Vol. 19 (2), pp. 537–544, 2004.

[7] McArthur S., et al. The application of model based reasoning within a decision support system for protection engineers, IEEE Trans. Power Delivery, Vol. 11 (4), pp. 1748–1754, 1996.

[8] Foong O., Sulaiman S., Rambli D., Abdullah N. ALAP: Alarm Prioritization System For Oil Refinery, Proc. of the World Congress on Engineering and Computer Science Vol II, San Francisco, CA, USA, 2009.

[9] Smarandache F., Dezert J. (Editors) Advances and Applications of DSmT for Information Fusion, ARP, Rehoboth, Vol.1-3, 2004-2009.

[10] Tchamova A., Dezert, J., Intelligent alarm classification based on DSmT, 6th IEEE International Conference "Intelligent Systems" Sofia, Bulgaria, pp.120-125, 2012.

[11] Tchamova A., Dezert J., Smarandache F.,A New Class of Fusion Rules based on T-Conorm and T-Norm Fuzzy Operators, Information and Security, An International Journal, Vol. 20, pp.77-93, 2006.

[12] Tchamova A., Dezert J., On the behavior of Dempster's rule of combination and the foundations of Dempster-Shafer Theory, 6th IEEE International Conference "Intelligent Systems" Sofia, Bulgaria, pp.108-113, 2012.

[13] Behar V., et al. STAP Approach for DOA Estimation using Microphone Arrays, Signal Proc. Workshop, Vilnius, SPIE Proc., Vol. 7745, 2010.

[14] Shafer, G. A Mathematical Theory of Evidence, Princeton Univ., 1976.

[15] Smets P., Kennes R. The transferable belief model, Artif. Intel., 66 (2), pp. 191-234, 1994.

[16] Mendel J., Fuzzy Logic Systems for Engineering: A Tutorial, Proc. of the IEEE, pp. 345–377, 1995.

URREF Reliability versus Credibility in Information Fusion (STANAG 2511)

Erik Blasch
Kathryn B. Laskey
Anne-Laure Jousselme

Valentina Dragos
Paulo C. G. Costa
Jean Dezert

Originally published as Blasch E., Laskey K.B., Jousselme A.-L., Dragos V., Costa P.C.G., Dezert J., *URREF reliability versus credibility in information fusion (STANAG 2511),* Proc. of Fusion 2013 Int. Conference on Information Fusion, Istanbul, Turkey, July 9-12, 2013, and reprinted with permission.

***Abstract*— *For many operational information fusion systems, both reliability and credibility are evaluation criteria for collected information. The Uncertainty Representation and Reasoning Evaluation Framework (URREF) is a comprehensive ontology that represents measures of uncertainty. URREF supports standards such as the NATO Standardization Agreement (STANAG) 2511, which incorporates categories of reliability and credibility. Reliability has traditionally been assessed for physical machines to support failure analysis. Source reliability of a human can also be assessed. Credibility is associated with a machine process or human assessment of collected evidence for information content. Other related constructs for URREF are data relevance and completeness. In this paper, we seek to develop a mathematical relation of weight of evidence using credibility and reliability as criteria for characterizing uncertainty in information fusion systems.*

Keywords: *Reliability, Credibility, URREF, PCR5, STANAG2511*

I. INTRODUCTION

Information fusion is based on uncertainty reduction; wherein the International Society of Information Fusion (ISIF) *Evaluation of Techniques of Uncertainty Reasoning Working Group* (ETURWG) has had numerous discussions on definitions of uncertainty. One example is the difference between reliability and credibility, which is called out in NATO STANAG 2511 [1]. To summarize these ETURWG discussions, we detail an analysis of credibility and reliability.

Information fusion consumers comprise users and machines of which the man-machine interface requires understanding of how data is collected, correlated, associated, fused, and reported. Simply stating an uncertainty representation of "confidence" is not complete. From URREF discussions [2]:

reliability relates to the source, and
credibility refers to the content reported.

There are scenarios in which reliability and credibility need to be differentiated. Examples of information fusion application areas include medical, legal, and military domains. A common theme is involvement of humans in aggregating information. In many situations, there is cause for concern about the reliability of the source that may or may not be providing an accurate and complete representation of credible information. In cases where there is a dispute (e.g., legal), the actors each seek their own interests and thus are asked a series of questions by their own and opposing representations to judge the veracity of their statements.

Weight of Evidence (WOE) is addressed in various fields (risk analysis, medical domain, police, legal, and information fusion). In addition to credibility and reliability, *Relevance* assesses how a given uncertainty representation is able to capture whether a given input is related to the problem that was the source of the data request. A final metric to consider is *completeness*, which reflects whether the totality of evidence is sufficient to address the question of interest. These criteria relate to high-level information fusion (HLIF) [3] systems that work at levels three and above of the Data Fusion Information Group (DFIG) model. For the URREF, we then seek a mathematical representation the weight of evidence:

$$\text{WOE} = f(\text{Reliability, Credibility, Relevance, Completeness}) \quad (1)$$

where f is an function to be defined with operations on how to combine such as a utility analysis.

Sect. II. provides related research and Sect. III overviews information fusion. Sect. IV discusses the weight of evidence including relevance and completeness. Sect. V describes the modeling of reliability and credibility with Sect. VI providing a simulation over evidence processing. Sect. VII provides discussion and conclusions.

II. BACKGROUND

There are many examples of ***reliability*** analysis for system components [4]. Typically, a reliability assessment is conducted on system parts to determine the operational life of each component over the entire collection of parts [5]. A reliability analysis can consist of many attributes such as survivability [6], timeliness, confidence, and throughput [7, 8]; however the most notable is time to failure [9]. Reliability is typically modeled as a continuous analysis of a part; however, a discrete analysis can conducted for the number of failures in a given period of time [10]. Real-time analysis requires information fusion between continuous and discrete analysis over new evidence [11], covariance analysis [12, 13], and resource analysis [14] to control sensors.

To assess the performance of sensors (and operators) requires analysis of the physical reliability of components. Data fusion can aid in fault detection [15], predictive diagnostics [16], situation awareness [17], and system performance. A model of

reliability includes time-dependent measures for operational lifetime analysis and controllability which are aspects of a data fusion performance analysis [18]. Use of multiple systems can aid in reducing failures through redundancy or system reconfiguration in response to failed sensors [19] for such applications as robotics [20, 21], risk analysis for situation awareness [22, 23], and cyber threats [24, 25, 26].

Time-dependent measures such as times between failures are appropriate for processes that operate over time to produce a stream of outputs; and failures can render the output stream unreliable. For systems that respond to discrete queries or produce alerts, such as human operators in a fusion center or pattern recognition systems, reliability is assessed through correspondence between outputs and the actual situation. The confusion matrix (CM) is a typical measure [27]. Reliability also relates to the opinions of observers [28].

Credibility To analyze credibility of evidence, we can use probabilistic or credibilistic frameworks such as Bayes, Dempster-Shafer, or following proportional conflict redistribution (PCR) principle, etc. [29, 30, 31]. Credibility of a hypothesis can be assessed through its prior probability or belief; and also through conflict: information is more credible when it does not conflict with other information.

To summarize,

- *Reliability* is an attribute of a sensor or other information source, and measures the consistency of a measure of some phenomenon. Reliability can be assessed by variance, probability of occurrence, and/or a small spatial variance of precision.
- *Credibility,* also known as believability, comprises the content of evidence captured be a sensor which includes *veracity*, *objectivity*, *observational sensitivity*, and *self-confidence.*

Reliability from the engineering design domain (e.g., mean time between failures) refers to consistent ability to perform a function, and reliability of a source means consistently measuring the target phenomenon. It may be useful to model source failures over time using an exponential or Poisson distribution. For information fusion and systems analysis, we need both a source element (reliability) as well as a content element (credibility) to characterize information quality. Next, we describe the information model that consists of data sources from human and machines that requires uncertainty analysis.

III. Information Fusion

A. Information Fusion Evaluation

Information fusion combines information from multiple sources, distributions [32], or information over various system-level model processing levels as described in the *Data Fusion Information Group* (DFIG) model [33, 34, 35], depicted in Figure 1. The DFIG model outlines various processes for information fusion such as object assessment [36] (Level 1 – L1), situational assessment (L2), impact assessment (L3), and resource management (L4). Data and information fusion can be applied to assess the operating performance of algorithms [37], sources (reliability), as well as message content (credibility). For system-level analysis, it is important to look at source context reliability of humans (L5) and data sources for sensor (L4) and mission management (L6).

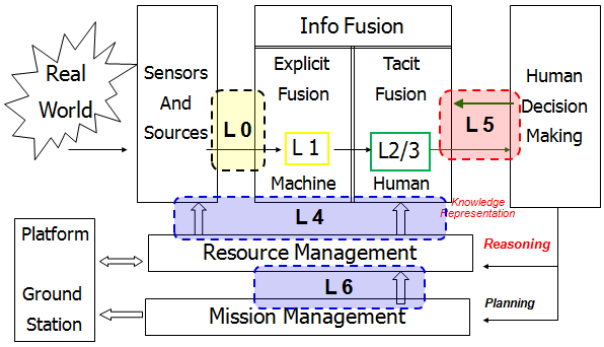


Figure 1 - DFIG Information Fusion model.

In the DFIG model, the goal is to separate information fusion (L0-L3) from sensor control, platform placement, and user selection to meet mission objectives (L4-L6) [38, 39, 40]. Information fusion across all the levels includes many metrics that need to be evaluated over uncertainty measures [41]. Challenges for information fusion, both at the hardware (i.e. components and sensors) and the software (i.e. algorithms and processes) levels were addressed by the *ETURWG* [http://eturwg.c4i.gmu.edu] [2]. Definitions of uncertainty measures such as accuracy [42], precision [43], reliability, and credibility are important for measures of effectiveness including validity and verification [44]. For example, accuracy (i.e., validity) measures distance from the truth, while precision (i.e., reliability) measures repeatability of results.

Examples of information fusion include tracking accuracy [45, 46], tracking filter credibility [47], and object detection credibility [48, 49] which are important for information quality and quality of service metrics [50].

B. NATO STANAG 2511

For STANAG 2511, as an update to STANAG2022, there are general listings of categories for reliability and credibility that are of interest to the ETRUWG [51, 52, 53]. Table 1 lists the STANAG 2511 issues that provided initial discussion for the ETURWG and the subsequent discussions in the URREF. Reliability and credibility are independent criteria for evaluation. The resultant rating will be expressed in the appropriate combination of letter and number (STANAG 2511). Thus information received from a "usually reliable" source which is adjusted as "probably true" will be rated as "B2". Information from the same source of which the "truth cannot be judged" will be rated as "B6".

The URREF ontology, shown in Figure 2, distinguishes between reliability and credibility in evidence handling and evidence processing; respectively. In this paper, we utilize the STANAG 2511 definitions of reliability (of source) and credibility (of information). From the ETURWG discussions, credibility and reliability also relate to weight of evidence, relevance, and completeness; although others are currently being explored.

Table 1: STANAG 2511 Reliability and Credibility Relations and Definitions

RELIABILITY	CODE	EXPLANATION From STANAG 2511
Completely Reliable	A	A tried and trusted source which can be depended upon with *confidence*
Usually Reliable	B	A past successful source for which there is still some element of doubt in particular cases
Fairly Reliable	C	A past occasionally used source upon which some degree of confidence can be based
Not Usually Reliable	D	A source which has been used in the past but has proved more often than not unreliable
Unreliable	E	A source which has been used in the past and has proved unworthy of any confidence
Cannot be judged	F	It refers to a source which has not been used in the past

CREDIBILITY	CODE	EXPLANATION From STANAG 2511
Confirmed	1	If it can be stated with *certainty* that the reported information originates from another source than the already existing information on the same object
Probably true	2	If the independence of the source cannot be guaranteed, but if, from the quantity and quality of previous reports, its *likelihood* is nevertheless regarded as sufficiently established
Possibly true	3	If insufficient confirmation to establish any higher degree of likelihood, a freshly reported item of information *does not conflict* with the previously reported target behavior
Doubtful	4	An item of information which tends to *conflict* with the previously reported or establish behavior pattern of an intelligence target
Improbable	5	An item of information which positively *contradicts* previously reported information of conflicts with the established behavior pattern of an intelligence target in a marked degree
Cannot be judged	6	If its truth cannot be judged

IV. Weight of Evidence

Weight of evidence (WOE) has different meanings in different contexts. A commonality is the need to integrate different sources or lines of evidence to form a conclusion or a decision.

In the field of *risk analysis*, WOE consists of a set of methods developed to assess the level of risks associated to factors or causes [54]. In most cases, WOE is a means of synthesizing information, while the solution adopted for weighing evidence is not explicit, or the evidence is presented without any interpretation. While some approaches rely on scoring techniques (see for instance research on sediments assessment described in [55]), the overall solutions remain qualitative in nature, developed for particular applications and poorly adaptable. Further discussion on WOE, as tackled within the risk analysis area is provided in [56].

WOE is addressed in a similar way in the *medical domain*, in relation to the rise of a new set of medical practices known as "evidence based medicine", promoting clinical solutions supported by practical experience, for which scientific support is not (yet) available.

From a different perspective, WOE is used in the *law and policy domain* to convey a subjective assessment of an expert analyzing different items of evidence, most often in relation to a causal hypothesis [57]. Intuitively, the concept is used to signify that the value of evidence must be above a critical threshold to support decisions or conclusions. In *law*, standards of evidence are recognized (for instance a three-level standard classifies evidence as "preponderance", "clear and convincing" and "beyond a reasonable doubt"), but experts will still exercise their judgment on the strength of evidence, as there is no methodology to assess this parameter. Without such methodologies, the variance in expert's judgments could be important, as subjective factors shape inevitably the outcome of the evidence in the evaluation.

A. Weight of evidence for information fusion

In the field of information fusion, WOE captures the intuition

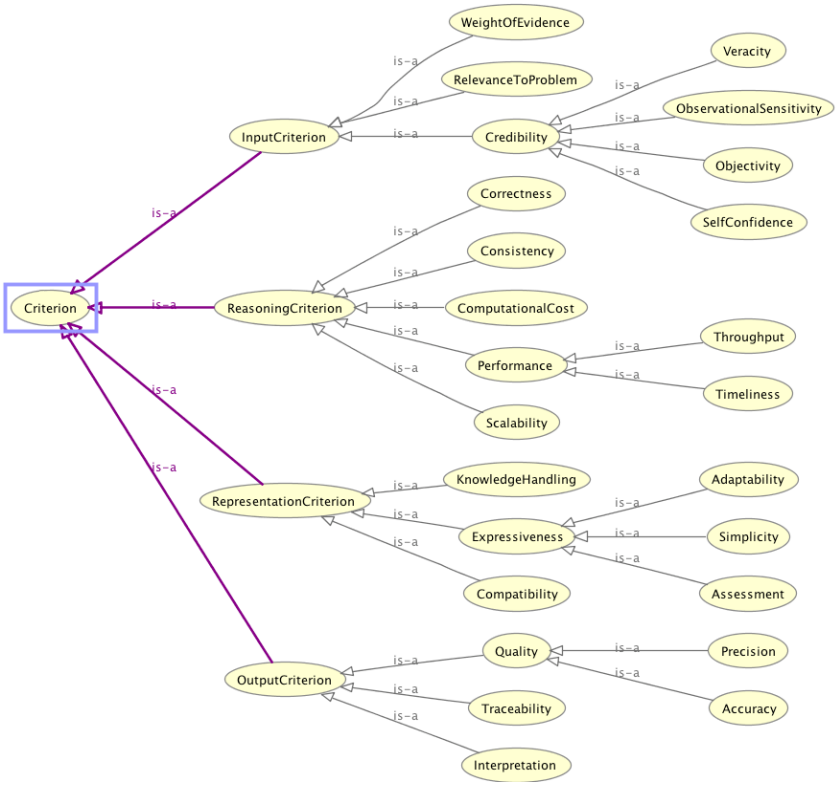


Figure 2 – URREF Ontology: Criteria Class [2].

that there is more or less evidence in the data, and this can be related to different parameters: the value of information itself (whether a piece of evidence conveys rich or poor information), the credibility of this information, in conjunction with the reliability of its source (can or should we believe this information), and finally the utility (or completeness) of this information with respect to a considered goal or task (is this data adding any detail to our existent data set?). WOE is an *attribute of information* and its values should be assessed by following a justifiable, repeatable and commonly accepted process. Therefore, several solutions have been developed to propose assessment mechanisms.

Among them, [58] proposes a probabilistic approach for information fusion where data items are weighted with respect to the accuracy or reliability of their source. This solution considers only independent information items and its adaptation to correlated information was developed [59]. In the field of evidential reasoning, the discounting operation introduced by Shafer [60], allows us to consider knowledge about the reliability of information sources. Smets and colleagues propose a method for learning a sensor's reliability, at various detail levels defined by users [61]. This method is generalized in Mercier, *et. al*. [62] by introducing the contextual discounting.

From a different perspective, [63] extends this frame in order to combine sources having different reliabilities and importance levels, while making a clear distinction between those notions.

It should be noticed that all references above consider only attributes of sources, while the weight of evidence should also be a function of information credibility. Underlying the same intuition of assigning different importance levels to items when fusing information, we can also cite research on prioritized and weighted aggregation operators, described in [64] and [65].

B. Relevance in Information Fusion

Relevance has these components: property relation and piece of evidence (POE). Relevance is often considered as a relation between one property (or feature) and a conditional. That means that a property is relevant (or related) to another one "if it leads us to change our mind concerning whether the second property holds" [66].

For instance, in classification, relevance criteria determine how well a feature (a property) discriminates between the classes (another property). In this case, the feature selection step aims at identifying the features that are most relevant to the classification problem. We distinguish between the filter mode and the wrapping mode. In the filter mode, measures of relevance are used to characterize the features. In the wrapping mode, a classifier is used and the optimal subset of relevant features is the one which maximizes the given performance measures, such as the recognition rate, the area under curve, etc., subject to a penalty on the number of features. Classical relevance measures are based on: mutual information, distances between probabilities, cardinality distances, etc.

A *piece of evidence* (POE) is relevant if it impacts previous beliefs. In this case, the relevance of a piece of information can only be evaluated in conjunction with the combination (updating, revision) operator used, as the null element and the properties in general may differ from one operator to another. For example, in Information Retrieval, the process is used to assess the relevance of retrieved items (documents) based on a given query.

Measures of relevance are based on traditional recall and precision measures: Precision is the fraction of retrieved items that are relevant, and Recall is the fraction of relevant items that have been retrieved [67].

Relevance is defined with respect to a goal (or a context) and assesses quantitative and qualitative information change.

- *Quantitative approaches*: In quantitative approaches, the notion of relevance is often intimately linked to the notion of independence. For instance, in classical probability theory, according to Gärdenfors [68], a proposition p is relevant to another proposition r on evidence e if p and r are conditionally dependent given e.
- *Qualitative approaches*: In qualitative approaches, the notion of relevance is linked to the material implication (see for instance the work of Goodman [69]): If a then b, $a \rightarrow b$, then a should be relevant to b.

C. How to evaluate a Relevance Criterion?

First, we should clarify what is the object under evaluation, or what do we mean by uncertainty representation (UR). We follow here the distinction put forward in [70] about the difference between uncertainty *calculi* and decision procedures.

If UR means uncertainty *calculus* (UC) (mathematical framework, theory), then we are asking if, for instance, possibility theory or probability theory is able "*to capture how a given input is relevant* [...]", and to what degree. Although this is a very general question with certainly no binary answer, some evaluation could be done.

For instance, using a literature survey for document retrieval, what is needed is a notional scale. An example of a scale to be defined over methods, measures, or models :

A. exist and are well developed with the theory and results are significant;
B. exist but some further developments are required or results are not significant;
C. are missing, or
D. the concept is not addressed.

We could conclude for instance probability theory is very good at dealing with relevance since a plethora of methods and measures are defined (A), compared to possibility theory for which only few methods exist (B). This would be an *empirical evaluation*, mainly based on a literature survey. Although we could conclude that a theory is very good at dealing with the relevance concept (numerous methods, measures, papers etc), an absence of evidence in this sense for another theory would not mean that the latter is not good. Rather it would identify a research gap.

Each of the following elements can be evaluated separately:

(UC-1) The mathematical model for uncertainty representation

(UC-2) The uncertainty measures
(UC-3) The inference rules and combination operator
(UC-4) Transformation functions

If UR is a *decision procedure* (DP), we are asking if a particular algorithm, relying on possibly several theories, is able "*to capture how a given input is relevant* [...]", and to what degree. A DP distinguishes between the method and its implementation (e.g., fusion algorithm). Also, note that the same DP could be represented by several algorithms.

Two steps underlying may be distinguished:

1. Identification and assessment of pieces of information (or properties) according to their relevance; and
2. Filtering of irrelevant pieces of information.

Example of an experiment to be elaborated could be:

i. Consider a dataset with both relevant and irrelevant pieces of information;
ii. Each piece of information should have been previously labeled as relevant or irrelevant, possibly with some degrees;
iii. Run the decision procedure (fusion algorithm) with only relevant pieces of information and add progressively irrelevant (or less relevant) ones; and
iv. Evaluate the decision procedure based on other independent criteria such as the execution time, true positive rate, conclusiveness, interpretation, etc.

We could observe for instance that a given Decision Procedure, say DP-A, is better than another one, say DP-B, because its execution time is lower with an equivalent true positive rate. Even if DP-A is based on evidence theory and DP-B is based on probability theory, concluding that evidence theory is better for dealing with relevance than probability theory is obviously not trivial and would require special care.

A thinner-grained assessment of relevance criterion can be performed by assessing separately each of the following elements of an Atomic Decision Procedure (ADP):

(ADP-1) Universe of discourse
(ADP-2) Instantiated uncertainty representation
(ADP-3) Reasoning step
(ADP-4) Decision step

For instance, one could assess if one particular universe of discourse better allows expressing *relevance* concepts than another. Relevance contributes to WOE. Evaluating whether a representation is able to deal with relevance should rely on other criteria of the ontology (if UR is a decision procedure) and or on other empirical criteria to be defined (if UR is an uncertainty calculus). In addition to relevance affecting reliability and credibility, completeness needs to be considered.

D. Evidence Completeness

Reliability versus credibility is highly related to completeness of evidence. For example, we cannot postulate that: (P1) reliability of a source => credibility of information (that is more a source is reliable, more the credibility of the information it provides is high) WITHOUT assuming the completeness of pieces of evidences available for the source.

For example: (*Ming vase*): Let's consider an apparent Ming vase (a counterfeit or a genuine one) to be analyzed. Suppose that an expert provides his report based on only two attributes (say the shape and color of the vase) and concludes (based on these two attributes/pieces of evidences only) that the vase is a genuine Ming vase. Because it is based on this knowledge only, and because both attributes fit perfectly with those of a genuine Ming vase, the Expert is 100% reliable (he didn't make a mistake) in assessing the two attributes; however, we are still unsure of his reliability in assessing whether the vase is genuine. Additional POE if available may be 100% reliable and support the opposite conclusion. For example, let's suppose that when looking at the vase we see the printed inscription "Made in Taiwan". So we are now sure that we are facing a counterfeit Ming vase.

So we see that the reliability and credibility notions are highly dependent on the underlying completeness of pieces of evidence and the relationship of the evidence to the conclusion of interest. In the Ming vase example, if we treat the two attributes (color and shape) as complete evidence sufficient to establish the absolute truth, then if Expert is fully reliable, the information he/she provides becomes highly credible due to reliability of the source and completeness of the evidence.

When there is incompleteness of POE, nothing conclusive can be inferred about credibility unless some additional assumptions are introduced about the evidence necessary to establish the truth.

The fundamental question behind this, is to know if a source based only on local/limited knowledge (evidences) can (or not) conclude with an absolute certainty about an hypothesis, or its contrary so that any other/additional pieces of evidences cannot revise his/her conclusion. Depending on the standpoint we choose, we accept or reject (P1) which makes a big difference in reasoning. In summary, the ETURWG analysis highlights uncertainty elements of a WOE.

E. URREF Weight of evidence

With respect to criteria defined by URREF we can define weight of evidence as:

$$\text{WOE} = f(\text{Reliability, Credibility, Relevance, Completeness})$$

where f is an function to be defined and relevance is related to the problem (or mission).

This is a translation of the following reasoning:

If (the source is reliable) then
If (the information provided is credible) then
If (this information is relevant to my problem) then
If (this information can enrich my existent information set) then
this information has some weight of evidence.

The four terms above are URREF criteria, while the last corresponds to a task-specific parameter that affects utility. For instance, utility can be evaluated by taking into consideration a distance between the set of information already available and a new item to determine utility completeness. Next, we demonstrate a modeling technique that brings together reliability and credibility to instantiate WOE calculations.

V. Reliability and Credibility Analysis

A reliability assessment affects modern equipment systems performance capability, maintainability, usability, and the operational support cost. Knowing the system's reliability is important for efficient and effective performance. Due to the high complexity of system's engineering integration, it is difficult to evaluate system-level reliability. Some ways to estimate system-level reliability include: (1) predicting operational reliability based on design data, (2) statistically analyze operational data, or (3) develop performance models based on real-world operational constraints.

Reliability prediction depends on models, such as life-cycle analysis. Typical models include Poisson, Exponential, Weibull, or Bernoulli distributions. Standard components, operating for a long time, may have data to support a priori analysis and modeling; however, the likelihood of reliability effectiveness is subject to real-world conditions that have not been modeled. For exponentially distributed failure times, the density function and the cumulative distribution function for time to failure of the system components are:

$$f(t) = \lambda e^{-\lambda t} \quad ; \quad F(t) = 1 - e^{-\lambda t} \qquad (2)$$

The physical meaning of $F(t)$ is the probability that a failure (doubt) occurs *before* the time t and $f(t)$ is the failure density: the probability that the component will fail in a small interval $t \pm \Delta t$ is given by $2f(t)\Delta t$. As t increases, the value of $F(t)$ approaches 1 at $t = \infty$.

For a fusion or reliability metric of a source, we need to map the semantics into quantifiable metrics based on the source context. Here we assume that we take discrete measurement and a consistent source has almost no failures. On the other hand, a non consistent source fails quickly. As a quick look we show a notional example, but realize that for human sources this model does not hold. For example, to ascertain a "not usual source" is difficult to quantify and caution and improvements would be forthcoming from the ETURWG.

Classification systems process evidence features by an algorithm to classify evidence into classes. Results are tested against truth and reported using a confusion matrix (CM) [27]. A CM can thus be used to measure reliability of a classification system. A CM is an estimate of likelihoods of the accumulated evidence of classifier. The elements of a confusion matrix are $c_{ij} = \Pr\{\text{Classifier decides } o_j \text{ when } o_i \text{ is true}\}$, where i is the true object class, j is the assigned object class, and $i = 1, \ldots, N$ for N true classes. The CM elements can be represented as probabilities as $c_{ij} = \Pr\{z = j \mid o_i\} = p\{z_j \mid o_i\}$. To determine an object declaration, we need to use Bayes' rule to obtain $p\{o_i \mid z_j\}$ which requires the class priors, $p\{o_i\}$. We denote the priors and likelihoods as column vectors

$$p(\overline{o}) = \begin{bmatrix} p(o_1) \\ p(o_2) \\ \vdots \\ p(o_N) \end{bmatrix} \; ; \; p(z_j \mid \overline{o}) = \begin{bmatrix} p(z_j \mid o_1) \\ p(z_j \mid o_2) \\ \vdots \\ p(z_j \mid o_N) \end{bmatrix}. \qquad (3)$$

For M decisions, a confusion matrix would be of the form

$$C = \begin{bmatrix} p(z_1 \mid o_1) & p(z_2 \mid o_1) & .. & p(z_M \mid o_1) \\ p(z_1 \mid o_2) & p(z_2 \mid o_2) & .. & p(z_M \mid o_2) \\ \ldots & \ldots & \ddots & \ldots \\ p(z_1 \mid o_N) & p(z_2 \mid o_N) & .. & p(z_M \mid o_N) \end{bmatrix}. \qquad (4)$$

VI. Results

For the simulation, we do both reliability and credibility assessment formulation to model the STANAG2511 criteria for uncertainty representation. Note that we assume completeness and relevance in these simulations.

A. Reliability

For source reliability, the parameter of choice is λ, which captures the rate of time between failures. Figure 3 demonstrates the intuition that reliable and unreliable sources remain unreliable and reliable. However, the interesting cases are those which are termed "usually reliable" (code B) which affects the uncertainty analysis.

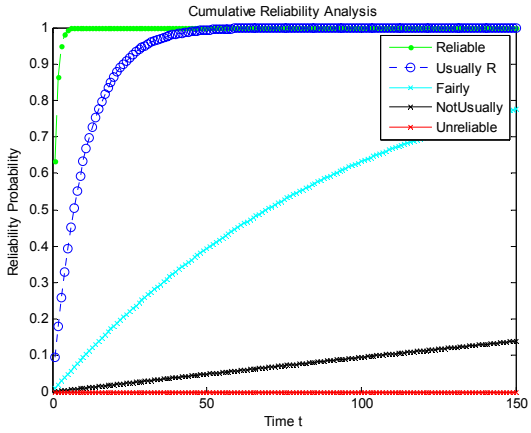


Figure 3 – Reliability Analysis

For Figure 3, a representative analysis of the reliability parameters are:

Code A	Completely Reliable :	λ = 0
Code B	Usually Reliable :	λ = 0.001
Code C	Fairly Reliable :	λ = 0.01
Code D	Not Usually Reliable :	λ = 0.1
Code E	Unreliable :	λ = 1
Code F	Cannot be judged	λ undefined

C. Credibility

For credibility, since STANAG 2511 definitions deal with conflicts, we utilize comparisons between Dempster-Shafer Theory and the PCR5 rule. Setting up the modeling using CM of classifiers from the information content, we can develop representative CMs for the different definitions:

```
%%% Confusion Matrices for Classifiers (two sources)
CM1=[0.999 0.001; 0.001 0.999]
CM2=[0.95 0.05; 0.05 0.95]
CM3=[0.70 0.30; 0.30 0.70]
```

Now, we define credibility levels as follows, based on the confusion matrices of the two classifiers and whether or not their outputs agree:

Confirmed:	CM1, outputs agree
Probably (independently confirmed):	CM2, outputs agree
Possibly (does not conflict):	CM3, outputs agree
Doubtful (tends to conflict):	CM3, outputs disagree
Improbable (conflicts):	CM1 or CM2, outputs disagree

Figure 4 shows a comparison of the CM results of a "possibly true" (code 3) to validate that the PCR5 rule better supports evidence analysis than the Dempster-Shafer method.

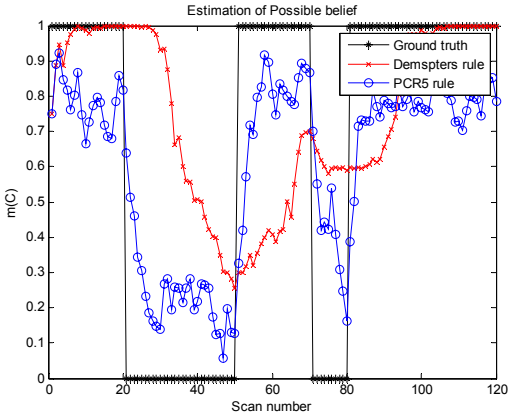


Figure 4 – DS versus PCR5 for "Possibly True" (Code 3)

Figure 5 and 6 highlight the credibility relations associated with a DS and PCR5 formulation, where PCR5 better represents an expected analysis for calculating the STANAG 2511 credibility codes.

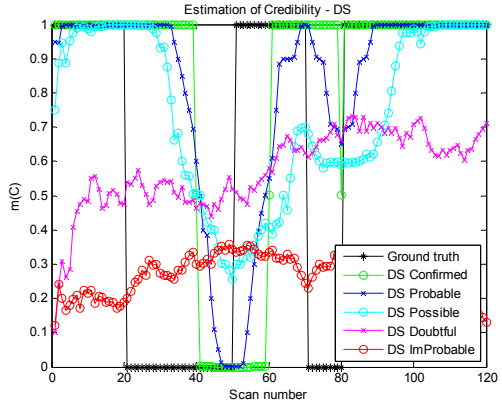


Figure 5 –DS Credibility of STANAG 2511 (Codes 1-5)

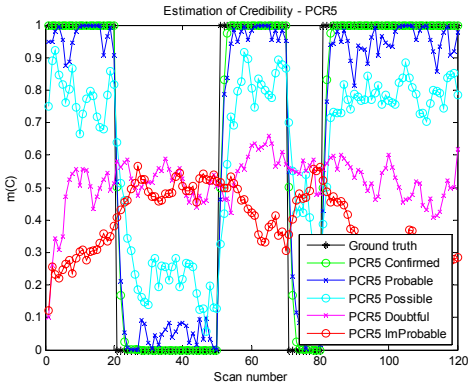


Figure 6 – PCR5 Credibility of STANAG 2511 (Codes 1-5)

VII. Conclusions

In this paper, we overviewed uncertainty representation discussions from the ETURWG as related to the STANG 2511 reliability and credibility. In our URREF model for weight of evidence, included are relevance and completeness. We demonstrated modeling for reliability and credibility and provided simulations as related to evidence reasoning methods of the PCR5 rule. These results provide a more tractable (and mathematical) ability to calculate the STANAG 2511 codes.

Reliability and credibility affect higher levels of information fusion (i.e. beyond Level 2 fusion) grand challenges [71] of uncertainty representation [72], ontologies [73, 74] and uncertainty evaluation [75, 76]. Future research will further explore the uncertainty ontology within the URREF, use cases of real systems for a combined credibility/reliability assessment, and mathematical inclusion of other metrics such as relevance and completeness.

References

[1] STANAG 2511 (January 2003, Intelligence reports, NATO Unclassified).

[2] P. C. G. Costa, K. B. Laskey, E. Blasch and A-L. Jousselme, "Towards Unbiased Evaluation of Uncertainty Reasoning: The URREF Ontology," *Int. Conf. on Info Fusion*, 2012.

[3] E. P. Blasch, E. Bosse, and D. Lambert, *High-Level Information Fusion Management and Systems Design*, Artech House, Norwood, MA, 2012.

[4] A. Peiravi, "Reliability Prediction of Electronic Navigation and Guidance Employing High Quality Parts to Achieve Increased Reliability," *J. Of App. Sciences*, 9 (16), 2009.

[5] R. Pan, "A Bayes Approach to Reliability Prediction Utilizing Data from Accelerated Life Tests and Field Failure Observations," *Quality and Reliability Engineering, International.* 25(2): 229-240, 2009.

[6] L. Bai, S. Biswas, and E. P. Blasch, "Survivability – An Information Fusion Process Metric From an Operational Perspective," *Int. Conf. on Info Fusion*, 2007.

[7] E. P. Blasch, M. Pribilski, B. Roscoe, *et. al.*, "Fusion metrics for dynamic situation analysis," *Proc. of SPIE,* Vol. 5429, 2004.

[8] E. Blasch, "Sensor, User, Mission (SUM) Resource Management and their interaction with Level 2/3 fusion" *Int. Conf. on Info Fusion*, 2006.

[9] E. P. Blasch, "Derivation of a Reliability Metric for Fused Data Decision Making," *IEEE NAECON Conf.*, 2008.

[10] L. Bai and E. P. Blasch, "Two-Way Handshaking Circular Sequential k-out-of-n Congestion System," *IEEE Trans. on Reliability*, Vol. 57, No. 1, pp. 59-70, Mar. 2008.

[11] E. Blasch, *Derivation of A Belief Filter for High Range Resolution Radar Simultaneous Target Tracking and Identification*, Ph.D. Dissertation, Wright State University, 1999.

[12] Y. Wu, J. Wang, J. Cheng, H. Lu, E. Blasch, L. Bai, and H. Ling, "Real-Time Probabilistic Covariance Tracking with Efficient Model Update," *IEEE Trans. on Image Processing*, **21**(5):2824-2837, 2012.

[13] C. Yang, L. Kaplan, and E. Blasch, "Performance Measures of Covariance and Information Matrices in Resource Management for Target State Estimation," *IEEE Trans. on Aerospace and Electronics*, Vol. 48, No. 3, pp. 2594 – 2613, 2012.

[14] E. Blasch, I. Kadar, K. Hintz, *et. al.*, "Resource Management Coordination with Level 2/3 Fusion Issues and Challenges," *IEEE Aerospace and Elect. Sys. Mag.*, Vol. 23, No. 3, pp. 32-46, Mar. 2008.

[15] M. Šimandl, M., and I. Punčochář, "Active-Fault detection and control: Unified formulation and optimal design," Vol. 45, *Automatica*, 2009.

[16] C. S. Byinton and A. Garga, "Data Fusion for Developing Predictive Diagnostics for Electromechanical Systems," Ch 23. in *Handbook of Data Fusion*, D. Hall and J Llinas (Eds.), CRC Press, 2001.

[17] E. Blasch, I. Kadar, J. Salerno, M. M. Kokar, S. Das, *et. al.*, "Issues and Challenges in Situation Assessment (Level 2 Fusion)," *J. of Advances in Information Fusion,* Vol. 1, No. 2, pp. 122 - 139, Dec. 2006.

[18] J. Llinas, "Assessing the Performance of Multisensor Fusion Processes," Ch 20 in *Handbook of Multisensor Data Fusion*, D. Hall and J. Llinas (Eds.), CRC Press, 2001.

[19] J. N. Yoo and G. Smith, "Reliability modeling for systems requiring mission reconfigurability," *Proc. Reliability and Maintainability Symp.*, pp. 133–139, 1990.

[20] G. Seetharaman, A. Lakhotia, E. Blasch, "Unmanned Vehicles Come of Age: The DARPA Grand Challenge," *IEEE Computer Society Magazine*, Vol. 39, No. 12, pp. 26-29, Dec 2006.

[21] K. M. Lee, Z. Zhi, R. Blenis, and E. P. Blasch, "Real-time vision-based tracking control of an unmanned vehicle," *Journal of Mechatronics - Intelligent Motion Control*, Vol. 5, No. 8, pp. 973 –991, 1995.

[22] T. Bass and R. Robichaux, "Defense – In-Depth Revisited: Qualitative Risk Analysis Methodology for Complex Network-centric operations," *IEEE MILCOM*, 2001.

[23] J. Salerno, E. Blasch, M. Hinman, and D. Boulware, "Evaluating algorithmic techniques in supporting situation awareness," *Proc. of SPIE,* Vol. 5813, April 2005.

[24] G. Chen, D. Shen, C. Kwan, J. Cruz, M. Kruger, and E. Blasch, "Game Theoretic Approach to Threat Prediction and Situation Awareness," *J. of Advances in Information Fusion,* Vol. 2, No. 1, 1-14, June 2007.

[25] G. Chen, E. P. Blasch, and L. Haynes, "A game Theoretic Data Fusion Approach for Cyber Situational Awareness," *Cyber Fusion Conf*., 2007.

[26] D. Shen, G. Chen, J. Cruz, et al., "Game Theoretic Solutions to Cyber Attack and Network Defense Problems," *ICCRTS*, Nov. 2007.

[27] B. Kahler and E. Blasch, "Decision-Level Fusion Performance Improvement from Enhanced HRR Radar Clutter Suppression," *J. of. Advances in Information Fusion,* Vol. 6, No. 2, Dec. 2011.

[28] J. Patel, *A Trust and Reputation Model for Agent-based Virtual Organizations*, PhD Univ. Southampton, 2007.

[29] J. Dezert, A. Tchamova, F. Smarandache, and P. Konstantinova, "Target Type Tracking with PCR5 and Dempster's Rules: a comparative analysis," *Int. Conf. on Information Fusion*, 2006.

[30] A. Martin, C. Osswald, J. Dezert, and F. Smarandache, "General Combination Rules for Qualitative and Quantitative Beliefs," *J. of Adv. in Information Fusion*, Vol. 3, No. 2, December, 2008.

[31] E. Blasch, J. Dezert, and P. Valin, "DSMT Applied to Seismic and Acoustic Sensor Fusion," *Proc. IEEE Nat. Aerospace Elect Conf*, 2011.

[32] E. P. Blasch and M. Hensel, "Fusion of Distributions for Radar Clutter modeling," *Int. Conf. on Info Fusion*, 2004.

[33] E. P. Blasch et al, "JDL Level 5 Fusion model 'user refinement' issues and applications in group Tracking," *Proc. SPIE,* Vol. 4729, 2002.

[34] E. P. Blasch, "Level 5 (User Refinement) issues supporting Information Fusion Management," *Int. Conf. on Info Fusion*, 2006.

[35] E. Blasch, *et. al*., "DFIG Level 5 (User Refinement) issues supporting Situational Assessment Reasoning," *Int. Conf. on Info Fusion,* 2005.

[36] C. Yang and E. Blasch, "Kalman Filtering with Nonlinear State Constraints," *IEEE Trans. Aero. and Elect. Systems*, Vol. 45, No. 1, 70-84, Jan. 2009.

[37] Z. Liu, E. Blasch, Z. Xue, R. Langaniere, and W. Wu, "Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, **34**(1):94-109. 2012.

[38] E. P. Blasch and P. Hanselman, "Information Fusion for Information Superiority," *IEEE Nat. Aerospace and Electronics Conference,* 2000.

[39] E. Blasch, "Situation, Impact, and User Refinement," *Proc. of SPIE,* Vol. 5096, April 2003.

[40] E. Blasch, "User refinement in Information Fusion", Chapter 19 in *Handbook of Multisensor Data Fusion 2nd Ed,* (Eds.) D. Hall, and J. Llinas, CRC Press, 2008.

[41] P. Hanselman, C. Lawrence, E. Fortunano, B. Tenney, and E. Blasch, "Dynamic Tactical Targeting," *Proc. of SPIE,* Vol. 5441, 2004.

[42] C. Yang and E. Blasch, "Pose Angular-Aiding for Maneuvering Target Tracking," *Int. Conf. on Info Fusion*, July 2005.

[43] E. P. Blasch, *et. al*., "Ontology Alignment using Relative Entropy for Semantic Uncertainty Analysis," *Proc. IEEE NAECON*, 2010.

[44] E. Blasch, P. Valin, E. Bossé, "Measures of Effectiveness for High-Level Fusion," *Int. Conference on Information Fusion*, 2010.

[45] X. Mei, H. Ling, Y. Wu, E. P. Blasch, and L. Bai, "Minimum Error Bounded Efficient L1 Tracker with Occlusion Detection," *IEEE Computer Vision and Pattern Recognition*, 2011.

[46] Y. Wu, E. Blasch, G. Chen, L. Bai, and H. Ling, "Multiple Source Data Fusion via Sparse Representation for Robust Visual Tracking," *Int. Conf. on Info Fusion*, 2011.

[47] E. Blasch, O. Straka, J. Duník, and M. Šimandl, "Multitarget Performance Analysis Using the Non-Credibility Index in the Nonlinear Estimation Framework (NEF) Toolbox," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON)*, 2010.

[48] E. Blasch, R. Breton, and P. Valin, "Information Fusion Measures of Effectiveness (MOE) for Decision Support," *Proc. SPIE* 8050, 2011.

[49] Y. Zheng, W. Dong, and E. Blasch, "Qualitative and quantitative comparisons of multispectral night vision colorization techniques," *Optical Engineering*, Vol. 51, Issues 8, Aug. 2012.

[50] C. Bisdikian, L. M. Kaplan, M. B. Srivastava, D. J. Thornley, D. Verma, and R. I. Young, "Building principles for a quality of information specification for sensor fusion," *Int. Conf. on Information Fusion*, 2009.

[51] V. Nimier, "Information Evaluation: a formalization of operational recommendations," *Int. Conf. on Information Fusion*, 2004.

[52] J. Besombes, and A. R. d'Allonnes, "An Extension of STANAG2022 for Information Scoring," *Int. Conf. on Info. Fusion*, 2008.

[53] E. P. Blasch, R. Breton, and P. Valin, "Information Fusion Measures of Effectiveness (MOE) for Decision Support," *Proc. SPIE* 8050, 2011.

[54] Linkov I, Welle P, Loney D, Tkachuk A, Canis L, Kim JB, Bridges T., "Use of multicriteria decision analysis to support weight of evidence evaluation," *Risk Anaysis*, August, 31(8):1211-25, 2011.

[55] P. Chapman, "A decision making framework for sediment assessment developed for the Great Lakes," *Human and Ecological Risk Assessment*, 8(7):1641–1655, 2002.

[56] D. L. Weed, "Weight of evidence: a review of concept and method." *Risk Analysis,* 25(6), 1545-1557, 2005.

[57] S. Krimsky, "The weight of scientific evidence in policy and law," *American Journal of Public Health,* 95 (S1), S129-S136, 2005.

[58] L. Peijun, and S. Benqin, "Land cover classification of multi-sensor images by decision fusion using weight of evidence model," International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences*, XXII ISPRS Congress, 2012.

[59] M. Deng, "A Conditional Dependence Adjusted Weights of Evidence Model. Natural Resources Research," 18(4), 249- 258, 2009.

[60] G. Shafer, *A mathematical theory of evidence*. Princeton University Press, Princeton, N.J., 1976.

[61] Z. Elouedi, K. Mellouli, and Ph. Smets, "Assessing sensor reliability for multisensor data fusion with the transferable belief model" *IEEE Tr. on Systems, Man and Cybernetics* B, volume 34, pages 782-787, 2004.

[62] D. Mercier, T. Denoeux, and M.H. Masson, "Refined sensor tuning in the belief function framework using contextual discounting," *Proc. of IPMU*, Vol II, pages 1443-1450, 2006.

[63] F. Smarandache, J. Dezert J.-M. Tacnet, "Fusion of sources of evidence with different importances and reliabilities, *Int. Conf. on Info Fusion* 2010.

[64] R. Yager, "Prioritized aggregation operators," *International Journal of Approximate Reasoning*, Vol. 48, Issue 1, pages 263-274, 2008.

[65] R. Yager, "Prioritized operators and their applications," *Proc. of the 6th IEEE International Conference 'Intelligent Systems'*, 2012.

[66] J. P. Delgrande, and F. J. Pelletier, "A Formal Analysis of Relevance," *Erkenntnis*, 49(2):137–173, 1998.

[67] P. Borlund, "The Concept of Relevance in IR", *Journal of the American Society for Information Sciences and Technology*, 54(10):913–925, 2003.

[68] P. Gärdenfors, "On the logic of relevance," *Synthese*, 37:351–367, 1978.

[69] N. Goodman. About. *Mind*, 70:1–24, 1961.

[70] A.-L. Jousselme, and P. Maupin, "A brief survey of comparative elements for uncertainty calculi and decision procedures assessment", *Panel discussion, Int. Conf,. on Information Fusion*, 2012.

[71] E. P. Blasch, D. A. Lambert, P. Valin, *et. al*., "High Level Information Fusion (HLIF) Survey of Models, Issues, and Grand Challenges," *IEEE Aerospace and Elect. Sys. Mag.,* Vol. 27, No. 8, Aug. 2012.

[72] P.C.G. Costa, R.N. Carvalho, K.B. Laskey, and C.Y. Park, "Evaluating Uncertainty Representation and Reasoning in HLF systems," *Int. Conference on Information Fusion*, 2011.

[73] E. Blasch, "Ontological Issues in Higher Levels of Information Fusion: User Refinement of the Fusion Process," *Int. Conf. on Info Fusion,* 2003.

[74] P.C.G. Costa, KC Chang, K.B. Laskey, T. Levitt, and W. Sun, "High-Level Fusion: Issues in Developing a Formal Theory," *Int. Conf. on Information Fusion*, 2010.

[75] E. Blasch, P. C. G. Costa, K. B. Laskey, D. Stampouli, *et al*., "Issues of Uncertainty Analysis in High-Level Information Fusion – Fusion2012 Panel Discussion," *Int. Conf. on Info Fusion*, 2012.

[76] P. C. G. Costa, E. P. Blasch, K. B. Laskey, S. Andler, J. Dezert, A-L. Jousselme, and G. Powell, "Uncertainty Evaluation: Current Status and Major Challenges – Fusion2012 Panel Discussion," *Int. Conf. on Info Fusion*, 2012.

Application of New Absolute and Relative Conditioning Rules in Threat Assessment

Ksawery Krenc
Florentin Smarandache

Originally published as Krenc K., Smarandache F., *Application of New Absolute and Relative Conditioning Rules in Threat Assessment,* Proc. of Fusion 2013 Int. Conf., Istanbul, Turkey, July 9-12, 2013, and reprinted with permission.

***Abstract*—This paper presents new absolute and relative conditioning rules as possible solution of multi-level conditioning in threat assessment problem.**

An example of application of these rules with respect to target observation threat model has been provided.

The paper also presents useful directions in order to manage the implemented multiple rules of conditioning in the real system.

I. Introduction

Contemporary Command & Control systems operate with multiple sensors in order to elaborate consistent and complete information required for decision making [1]. These systems, however, must face another very important requirement, which is cooperation with other information systems. Dealing with information of different processing levels is inevitable consequence of the imposed demands, and requires specific tools for fusion in order to take this diversity into account effectively.

As Threat Assessment is one of the most important tasks imposed on C2 systems [2], [3], these systems must be able to deal with information obtained from uncertain and even unreliable sources, where the quality measures are often subjective. For this reason Theory of Evidence seems to be an appropriate approach.

Theory of Evidence known as Dempster-Shafer Theory (*DST*) [4] does not make any distinction to fusion operations regarding uncertainty of the gathered information. So called Dempster's rule of combination has been used in order to combine strong evidences from reliable sources, as well as poor evidences from unreliable sources, and hybrid (strong evidence with poor evidence). For many years researchers have been inventing diverse combination rules as alternative to Dempster's rule [5], [6], and [7]. These rules are different from each other mainly in the way the conflicting mass (referring to contradicting hypotheses) is distributed. However, according to knowledge of the authors, none of these rules takes into account possible different processing levels of the integrated information, which cannot be expressed with *basic belief assignments*.

Theory of Evidence by Dezert and Smarandache (*DSmT*) [8] distinguishes two operations: *combination* and *conditioning* for fusion of uncertain information and integration of uncertain pieces of information with confirmed i.e. certain evidence respectively. Aware of this fact, a certain idea of using *conditioning* operation (as an alternative of *combination* [9]) for the purpose of multiple level fusion has been published [10]. However, as it was presented in [11], [12], each of these solutions has its drawbacks, and in general neither is preferable over the other. The main disadvantage of *combination* as multiple level fusion operation is that it does not take into account the predominance of the conditioning information from the external system over the local sensor data, and in result it makes no distinction between the information processing levels. On the other hand, the main disadvantage of *conditioning* is that the condition is treated, by definition, as an absolute and literate fact, which is the assumption very hardly accepted in the real world.

For this reason another class of fusion rules, called *relative conditioning*, has been invented. In this type of rules the predominance of the condition over the uncertain evidence is stated explicitly, while the trust in the conditioning hypothesis is not absolute by definition.

In this paper two of these rules will be presented as possible solution of the multi-level conditioning [12] in threat assessment problem.

II. New Conditioning Rules

Let Θ be a frame of discernment formed by *n* singletons defined as:

$$\Theta = \{\theta_1, \theta_1, ..., \theta_n\}, n \geq 2 \tag{1}$$

and its Super-Power Set (or fusion space):

$$S^{\Theta} = (\Theta, \cup, \cap, C) \tag{2}$$

which means the set Θ is closed under union $\cup$, intersection $\cap$, and complement *C* respectively.

Let *m(.)* be a mass:

$$m(.) : S^{\Theta} \rightarrow [0, 1] \tag{3}$$

and a non-empty set A $\subseteq I_t$ where $I_t = \theta_1 \cup \theta_2 \cup ...\theta n$ is the total ignorance.

Conditioning of $m(.|.)$ becomes:

$$\forall X \in S^{\Theta}, m(X|A) = \sum_{\substack{Y \in S^{\Theta} \setminus \emptyset \\ Y \cap A = X}} m(Y) + \\ + \sum_{\substack{Y \in S^{\Theta} \setminus \emptyset \\ Y \cap A = \emptyset \\ X = A}} m(Y) \cdot \omega_A + \delta_X^A \cdot m(X) \cdot \omega_0 \quad (4)$$

where:

$$\delta_X^A = \begin{cases} 1, A = B \\ 0, A \neq B \end{cases} \quad (5)$$

and ω_0 and ω_A are the weights for all sets which are completely outside of A, and respectively for all sets which are inside or on the frontier of A.

$$\omega_0, \omega_A \in [0, 1], \omega_0 + \omega_A = 1 \quad (6)$$

For a more refined/ optimistic redistribution, all masses of the elements situated outside of A are redistributed, according to the formula (7).

$$\forall X \in S^{\Theta}, m(X|A) = \sum_{\substack{Y \in S^{\Theta} \setminus \emptyset \\ Y \cap A = X}} m(Y) + \\ + \frac{m(X)}{\sum_{\substack{Y \in S^{\Theta} \setminus \emptyset \\ Y \subseteq A \\ m(Y) \neq 0}} m(Y)} \sum_{\substack{Y \in S^{\Theta} \setminus \emptyset \\ Y \cap A = \emptyset \\ X = A}} m(Y) \cdot \omega_A + \delta_X^A \cdot m(X) \cdot \omega_0 \quad (7)$$

From the practitioner's point of view these formulas provide directions on how the mass of hypotheses not involved or partially involved in condition should be redistributed. In order to explain the idea of these rules it is suggested to consider a simple example of a model consisting of three hypotheses: A, B, and C, where A and B overlap each other, and C is disjoint. Assume the condition is A.

For this example, application of the rule (4) will cause the following action:

- former masses of A and A∩B remain unchanged, supplying A and A∩B hypotheses respectively,
- former mass of B is transferred to A∩B,
- former mass of C is transferred to A.

When applying the absolute version of the rule (4) all masses are transferred exactly as described above. Otherwise, i.e. relative conditioning, the mass of C is weighted according to the given ω_0 and ω_A.

Application of the rule (7) will cause the following action:

- former masses of A and A∩B remain unchanged, supplying A and A∩B hypotheses respectively,
- former mass of B is transferred to A∩B,
- former mass of C is transferred to A and A∩B proportionally to their masses

Similarly as for the rule (4) when applying the absolute version of the rule (7) all masses are transferred exactly as described above. Otherwise, i.e. relative conditioning, the mass C is weighted according to the given ω_0 and ω_A.

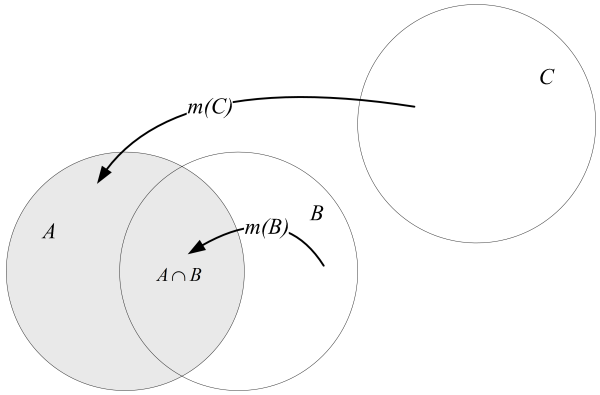


Figure 1. Mass transfer in case of application of the rule (4)

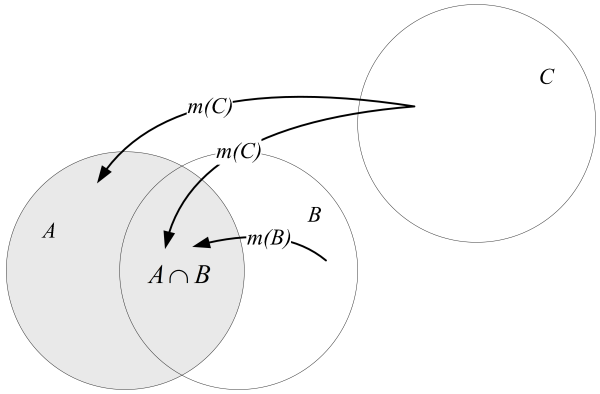


Figure 2. Mass transfer in case of application of the rule (7)

III. Threat Assessment Example

In order to illustrate application of the introduced rules it is suggested to consider the following conditioning example referring to the threat assessment problem. Assume the frame of discernment is defined as:

$$\Theta = \{F, H, U, N\} \quad (8)$$

where:

- F denotes FRIEND,
- H denotes HOSTILE,
- U denotes UNKNOWN,
- N denotes NEUTRAL.

Additionally assume:

$$S = H \cap U \quad (9)$$

$$A = F \cap U \quad (10)$$

$$K = F \cap H \quad (11)$$

$$J = F \cap H \cap U \quad (12)$$

where:

- S denotes SUSPECT,
- A denotes ASSUMED FRIEND,
- K denotes FAKER i.e. FRIEND acting as HOSTILE for training purposes, [13], [14], [15], and [16]

– J denotes JOKER i.e. FRIEND acting as SUSPECT for training purposes, [13], [14], [15], and [16].

Consider a scenario, where a local system, equipped with sensors and performing target threat observation and information fusion, gets informed by an external system about its decision, referring to the observed target. The decision transferred to the local system is that the target is FRIEND, which performs a conditioning information.

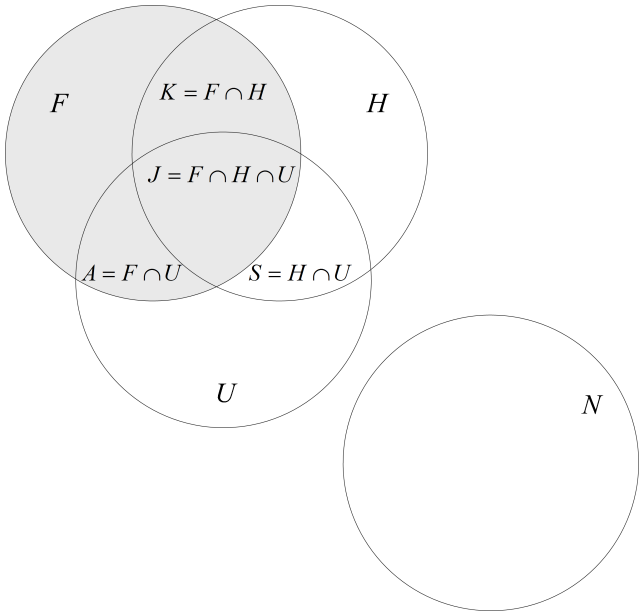


Figure 3. Venn's diagram of the observed target threat

Figure 3. shows a Venn's diagram describing the target threat observation model, where information obtained from the external system has been colored in gray. Notice that the model refers to observation of the target threat (not to the target threat in itself), which means it describes what the target looks like (not what the target really is). This is significant for justification why FAKER may be defined as the intersection of FRIEND and HOSTILE, not as a subset of FRIEND, which is by definition of FAKER in [13], [14], [15], and [16].

Consider that the local system has already performed sensor fusion and its results are summarized in basic belief assignment (*bba*) below:

$$\begin{array}{lll} m(F) = 0.2, & m(H) = 0.1, & m(U) = 0.1 \\ m(A) = 0.1, & m(S) = 0.1, & m(K) = 0.1 \\ m(J) = 0.2, & m(N) = 0.1 & \end{array}$$

Application of (4) leads to the following updated *bba* for absolute ($\omega_0 = 0$ and $\omega_A = 1$) conditioning:

$$\begin{array}{lll} m(F|F) = 0.3, & m(H|F) = 0, & m(U|F) = 0 \\ m(A|F) = 0.2, & m(S|F) = 0, & m(K|F) = 0.2 \\ m(J|F) = 0.3, & m(N|F) = 0 & \end{array}$$

For the relative conditioning with the following weights $\omega_0 = 0.3$ and $\omega_A = 0.7$ one should get:

$$\begin{array}{lll} m(F|F) = 0.27, & m(H|F) = 0, & m(U|F) = 0 \\ m(A|F) = 0.2, & m(S|F) = 0, & m(K|F) = 0.2 \\ m(J|F) = 0.3, & m(N|F) = 0.03 & \end{array}$$

For the absolute opposite ($\omega_0 = 1$ and $\omega_A = 0$) conditioning one should get:

$$\begin{array}{lll} m(F|F) = 0.2, & m(H|F) = 0, & m(U|F) = 0 \\ m(A|F) = 0.2, & m(S|F) = 0, & m(K|F) = 0.2 \\ m(J|F) = 0.3, & m(N|F) = 0.1 & \end{array}$$

Application of (7) leads to the following updated bba for absolute conditioning:

$$\begin{array}{lll} m(F|F) = 0.233, & m(H|F) = 0, & m(U|F) = 0 \\ m(A|F) = 0.217, & m(S|F) = 0, & m(K|F) = 0.217 \\ m(J|F) = 0.333, & m(N|F) = 0 & \end{array}$$

For the relative conditioning with the following weights $\omega_0 = 0.3$ and $\omega_A = 0.7$ one should get

$$\begin{array}{lll} m(F|F) = 0.223, & m(H|F) = 0, & m(U|F) = 0 \\ m(A|F) = 0.212, & m(S|F) = 0, & m(K|F) = 0.212 \\ m(J|F) = 0.323, & m(N|F) = 0.03 & \end{array}$$

For the absolute opposite conditioning one should get:

$$\begin{array}{lll} m(F|F) = 0.2, & m(H|F) = 0, & m(U|F) = 0 \\ m(A|F) = 0.2, & m(S|F) = 0, & m(K|F) = 0.2 \\ m(J|F) = 0.3, & m(N|F) = 0.1 & \end{array}$$

Analysis of the obtained results shows that there are substantial differences in results between conditioning rules (4) and (7) for the considered case. Depending on the particular rate of belief (values of ω_0 and ω_A) in condition the mass of the condition (FRIEND), as well as subsequent masses of hypotheses contained in the hypothesis of the condition (FAKER, JOKER, ASSUMED FRIEND) have been supplied with masses of hypotheses not contained in the condition (HOSTILE, UNKNOWN, SUSPECT, and NEUTRAL).

For both of the rules, in the first place the absolute conditioning case has been considered as a specific circumstance of relative conditioning. As the second, the relative conditioning has been performed with given weights of ω_0 and ω_A. Then, the absolute opposite conditioning has been presented as another special circumstance of relative conditioning.

The reason for the absolute opposite conditioning in this case is purely illustrative. Theoretically, it could be useful if the condition hypothesis was complex (expressed as union or intersection of multiple hypotheses) and it was convenient to consider the complement of the condition. However, in most of the cases the condition, as output of the external system is simple. Thus, it is very unlikely that such kind of conditioning would be applied in threat assessment.

Regarding the distinction in the presented rules, in this case, the essential difference between conditioning rules (4) and (7) resides in the manner the mass of NEUTRAL hypothesis is redistributed. For the rule (4) the mass of NEUTRAL is transferred completely to the mass of FRIEND, while for

the rule (7) the mass of NEUTRAL is transferred to FRIEND, JOKER, FAKER, and ASSUMED FRIEND proportionally to their masses. In other words, in case of the rule (7) the redistribution is performed with the higher degree of trust in the adequacy of the target threat observation model. Therefore it may be regarded as more optimistic in comparison to pessimistic rule (4).

IV. Choosing the proper conditioning rule

Choosing the proper rule is one of the most important questions related to application of any fusion techniques (conditioning and combination). Since there are many rules of combination and conditioning [8], [11], [9], [10], and even more possible fusion cases, the choice of any particular rule for the particular case could be a topic of papers for the next few decades. Moreover, since there are no existent standardized fusion cases for particular domains the choice of the optimal rule seems to be a philosophical problem.

Since in this paper there are two rules of conditioning proposed the problem of selection of the proper one still holds. Additionally, each of these rules introduces weights (ω_0 and ω_A) in order to establish the 'relativity' of the conditioning, and setting particular values to these weights requires a comment.

According to the knowledge of the authors [11], [9], [10], and [12], in most of the cases selection of the particular rule for conditioning (as well as combination) is done experimentally. For the particular fusion task e.g. threat assessment in Command and Control system one chooses the rule which returns the closest results to the expected values. However, even within the particular fusion task it is possible to find situations, where another rule returns results substantially better than the previously selected one. That means two things:

- there is no universal rule of conditioning, correct in every conditions,
- if that is so, the particular fusion task should be split for at least two subtasks.

In other words, the particular rule of conditioning should be selected dynamically according to specified circumstances of information integration process.

In this section, the authors would like to define the factors which may influence on the choice of the particular rule of conditioning.

Quality of gathered information could be regarded as a basic parameter that affects selection of conditioning rules. Further, this parameter may be decomposed for two components referring to attribute (observation) model and data. Thus, the quality aggregates both: model adequacy and data precision. The fundamental question is how these model adequacy and data precision may be assessed and transformed into the quality in order to make choice of conditioning rule?

Possible solution of this problem may reside in analysis of *bba* subjected to conditioning. *Bba*, by definition, performs a kind of distribution, where subsequent masses reflect the degree of belief in particular hypotheses. If sensors are not reliable relatively high mass will be transferred to hypothesis describing complete ignorance. For instance, for the considered case it could be I = F $\cup$ H $\cup$ U $\cup$ N. By implication if the sensors are reliable the mass referring to the complete ignorance is zero. That may be regarded as the first insight in data precision. Another inference on data precision may be done by overview of distribution of mass over the rest of the hypotheses. Conciseness of the distribution means higher precision. Adequacy of the attribute (observation) model, on the other hand, may be defined by compliance of hypothesis of the highest mass with the hypothesis of the condition. If there exists any relation between the highest mass hypothesis and the condition, e.g. including or intersecting they may be regarded as compliant. On the other hand if they are disjoint they are regarded as noncompliant.

Referring the deductions above to the features of the presented rules a simple logic (briefly described in Table I) may be applied in order to choose the proper conditioning rule.

Table I
Choice of the conditioning rule based on model adequacy and data reliability

Model	Data	Quality	Description	Conditioning
poor	poor	poor	$m_{max} \neq Cond$, $m(\Theta) \uparrow$	absolute, (4)
poor	good	poor	$m_{max} \neq Cond$, $m(\Theta) \downarrow$	absolute, (4)
good	poor	poor	$m_{max} \cong Cond$, $m(\Theta) \uparrow$	absolute, (7)
good	good	good	$m_{max} \cong Cond$, $m(\Theta) \downarrow$	relative, (7)

If the highest mass hypothesis is not compliant with the condition, which means the attribute (observation) model is not adequate, no matter if the data are precise or not, in such case absolute conditioning should be applied with no respect to the attribute (observation) model. This may be achieved by using the rule (4) with $\omega_0 = 0$ and $\omega_A = 1$.

If the mass referring to total ignorance is relatively high and the highest mass hypothesis is compliant with the condition that means that the sensor data are poor and the attribute (observation) model is adequate. In such case absolute conditioning should be applied with respect to the attribute (observation) model which may be achieved by using the rule (7) with $\omega_0 = 0$ and $\omega_A = 1$.

Finally, if the mass referring to total ignorance is relatively low and the highest mass hypothesis is compliant with the condition that means that the sensor data are reliable (good) and the attribute (observation) model is adequate. In such case relative conditioning should be applied with respect to the attribute (observation) model which may be achieved by using the rule (7) with $\omega_0, \omega_A \in (0, 1)$, where: $\omega_0 + \omega_A = 1$.

As a summary of this section it is worth of notice that particular values of the 'relativity' weights (ω_0 and ω_A) depend only on the specific configuration of the fusion system. In the authors' opinion it is pointelss to discuss any specific values without reference to the particular system since there are no general guidelines for presetting.

V. Selection of conditioning rules - examples

In order to illustrate the selection mechanism few more examples have been delivered. However, in the first place, it is

suggested to reconsider the example from Section III. Table II presents the summarized *bba* (before and after conditioning). In this very case, before conditioning performed, the dominant masses had referred to FRIEND and FAKER hypotheses, which was compliant with the condition hypothesis (FRIEND). That means the model was adequate. Additionally, the total ignorance mass has not been defined (as nonzero), which means the data were reliable. According to Table I, in such case the relative version of the rule (7) should be selected, which was exactly what was decided.

Table II
EXAMPLE 1: *BBA* BEFORE AND AFTER CONDITIONING OPERATION

Threat \ bba	m	$m_{(7)R}(\cdot\|F)$
F	0.2	**0.223**
H	0.1	0
U	0.1	0
A = F ∩ U	0.1	0.212
S = H ∩ U	0.1	0
K = F ∩ H	0.1	0.212
J = F ∩ H ∩ U	0.2	0.323
N	0.1	0.03
I = F ∪ H ∪ U ∪ N	0	0

In the next example it is suggested to consider *bba* given in the second column of the Table III. In this case the biggest mass has been assigned to total ignorance. Furthermore, there is no predominance of any particular primary hypotheses [17] (FRIEND, HOSTILE, UNKNOWN, NEUTRAL) or secondary hypotheses [17] (ASSUMED FRIEND, SUSPECT, FAKER, JOKER). That means that the gathered data are not reliable and and the model adequacy has not been proven. Therefore the absolute version of the rule (4) should be chosen.

Table III
EXAMPLE 2: *BBA* BEFORE AND AFTER CONDITIONING OPERATION

Threat \ bba	m	$m_{(4)A}(\cdot\|F)$
F	0.1	**0.56**
H	0.1	0
U	0.1	0
A = F ∩ U	0.06	0.16
S = H ∩ U	0.06	0
K = F ∩ H	0.06	0.16
J = F ∩ H ∩ U	0.06	0.12
N	0.06	0
I = F ∪ H ∪ U ∪ N	0.4	0

In the last example it is suggested to consider *bba* given in the second column of the Table IV. In this case the biggest mass has also been assigned to total ignorance, which proves relatively low sensor reliability. However, except *m(I)*, there is a predominance of FRIEND hypothesis over the other hypotheses. Thus the model be regarded as adequate. Therefore the absolute version of the rule (7) should be chosen.

In the presented procedure of selection of the conditioning rules *bba* provides qualitative information on data reliability as well as model adequacy. Analyzing the above examples, some harsh reader could regard reasoning about the adequacy of the model based on the *bba* as vague, due to the fact *bba*s are affected with measuring errors, and it is possible these errors influence on the decision whether a particular model is adequate or not. However, it is important to notice that in real systems these *bba*s are updated regularly, which enables to improve statistically the reference for decision making. That means that any predominance of a certain hypothesis may be confirmed by the subsequent version of updated *bba*.

Table IV
EXAMPLE 3: *BBA* BEFORE AND AFTER CONDITIONING OPERATION

Threat \ bba	m	$m_{(7)A}(\cdot\|F)$
F	0.16	**0.422**
H	0.1	0
U	0.1	0
A = F ∩ U	0	0.1
S = H ∩ U	0.06	0
K = F ∩ H	0.06	0.259
J = F ∩ H ∩ U	0.06	0.219
N	0.06	0
I = F ∪ H ∪ U ∪ N	0.4	0

It is also a matter of convention how to deal with a particular case when $m(\Theta) = m(I)$ is the maximal mass in the *bba*. Assuming that the condition hypothesis does not refer to total ignorance: On one hand, since *bba* influences both data reliability and model adequacy it is justified to select the absolute version of the rule (4). On the other hand, it is reasonable to exclude the total ignorance hypothesis $m(\Theta) = m(I)$ while deciding about the adequacy of the model, in order to distinguish two aspects (qualitative features) of the gathered *bba*, which is preferable by the authors.

VI. CONCLUSION

The introduced new rules of conditioning have been invented as a response for problems emerging while applying the existing absolute conditioning techniques in the real world. Considering the condition as identical with the ground truth may be useful in theory, however in practice it often performs an assumption hard to accept [11]. Updating attribute fusion results with evidence from the external system is an excellent example for that. Each time the highly processed information is used, no matter how good the system is, there is a risk that the output information is corrupted or at least slightly changed [18], [19].

The presented conditioning rules enable to set weights in order to define the degree of belief in the external system output. These weights should be treated as tactical and technical parameters of the system performing combination and conditioning. Certainly, depending on the actual needs, they may be fixed or changeable dynamically. However the exact values should result from the particular system configuration thus no theoretical preference is made.

In case of choosing a particular rule of conditioning it is different, and some general guidelines may be established. The proposed method of selection of the conditioning rules may be applied in Command and Control systems, where multiple rules may be implemented. In such case the choice of the proper conditioning rule may perform an element of so called Conditioning Management.

REFERENCES

[1] K. Krenc, A. Kawalec: *An evaluation of the attribute information for the purpose of DSmT fusion in C&C systems*, Fusion2008, Cologne, ISBN 978-3-00-024883-2, 2008.

[2] G. Hallingstad, K. Wrona: *Multisensor Cyber Defence Data Fusion*, Military Communications and Information Technology: A Comprehensive Approach Enabler, ISBN 978-83-62954-20-9, 2011.

[3] R. Piotrowski, B. Jasiul, M. liwka, G. Kantyka, T. Podlasek, T. Dalecki, M. Chora, R. Kozik, J. Brzostek: *The Response to Cyber Threats in Federation of Systems Environment*, Military Communications and Information Technology: A Comprehensive Approach Enabler, ISBN 978-83-62954-20-9, 2011.

[4] G. Shafer: *A Mathematical Theory of Evidence*,Princeton Univ. Press, Princeton, NJ, 1976.

[5] R. Yager: *On the Dempster-Shafer framework and new combination rules*, Information Science Vol. 41, pp 93-138, 1987.

[6] D. Dubois, H. Prade: *Representation and combination of uncertainty with belief functions and possibility measures*,Computational Intelligence, 4, pp. 244-264, 1988.

[7] Ph. Smets: *The combination of evidence in the transferable belief model*, IEEE-Pattern analysis and Machine Intelligence, Vol. 12, pp 447-458, 1990.

[8] F. Smarandache, J. Dezert: *Advances and Applications of DSmT for Information Fusion, Vol 1-3*,American Research Press Rehoboth, 2004-2009.

[9] K. Krenc, *Examination of information fusion performance for diverse Matlab implemented combination rules*, PB-11-09063-121, 2011.

[10] K. Krenc, *Examination of information fusion performance for diverse Matlab implemented conditioning rules*, PB-11-09063-122, 2011.

[11] K. Krenc, *Updating Attribute Fusion Results with Additional Evidence Using DSmT*, Fusion2012, Singapore, ISBN: 978-0-9824438-4-2, 2012.

[12] K. Krenc, *A frame of information fusion subsystem ontology for the purpose of C2 systems*, PB-11-09063-323, 2011.

[13] *The Joint C3 Information Exchange Data Model, Edition 3.1b*, 2007

[14] NATO Standardization Agreement: *Identification Data Combining Process, STANAG No. 4162*,Ed. 2.

[15] NATO Standardization Agreement: *Tactical Data Exchange - Link 16, STANAG No. 5516*,Ed. 3.

[16] NATO Standardization Agreement: *Tactical Data Exchange - Link 11, STANAG No. 5511*,Ed. 6.

[17] K. Krenc, A. Kawalec: *An application of DSmT in ontology-based fusion systems*, Fusion2009, Seattle, ISBN 978-0-9824438-0-4, 2009.

[18] A. K. Hyder et al. (eds.): *Multisensor Fusion*, ISBN 1-4020-0722-1, 99-124, 2002.

[19] K. Krenc, A. Kawalec, T. Pietkiewicz: *Definition of bba and quality of fusion in C2 systems*, Fusion2011, Chicago, ISBN: 978-0-9824438-2-8, 2011.

A DSmT Based Combination Systems for Handwritten Signature Verification

Nassim Abbas
Youcef Chibani
Bilal Hadjadji

Compilation of: N. Abbas and Y. Chibani, "SVM-DSmT Combination for Off-Line Signature Verification," IEEE International Conference on Computer, Information and Telecommunication Systems (CITS), Amman, Jordan, pp. 1-5, May 14-16, 2012., and of N. Abbas and Y. Chibani, "SVM-DSmT combination for the simultaneous verification of off-line and on-line handwritten signatures," International Journal of Computational Intelligence and Applications (IJCIA), vol. 11, no. 3, 2012, and reprinted with permission.

Abstract: *The identification or authentication from the handwritten signature is the most accepted biometric modality for identifying a person. However, a single handwritten signature verification (HSV) system does not allow achieving the required performances. Therefore, rather than trying to optimize a single HSV system by choosing the best features or classifier for a given system, researchers found more interesting to combine different systems. In that case, the DSmT is reported as very useful and powerful theoretical tool for enhancing the performance of multimodal biometric systems. Hence, we propose in this chapter a study of applying the DSmT for combining different HSV systems. Two cases are addressed for validating the effective use of the DSmT. The first one is to enhance the performance of off-line HSV systems by associating features based on Radon and Ridgelet transforms for each individual system. The second one is associating off-line image and dynamic information in order to improve the performance of single-source biometric systems and ensure greater security. Experimental results conducted on standard datasets show the effective use of the proposed DSmT based combination for improving the verification accuracy comparatively to individual systems.*

1.1 Introduction

Biometrics is one of the most widely used approaches for identification and authentication of persons [1]. Hence, several biometric modalities have been proposed in the last decades, which are based on physiological and behavioral characteristics depending on their nature. Physiological characteristics are related to anatomical properties of a person, and include for instance fingerprint, face, iris and hand geometry. Behavioral characteristics refer to how a person performs an action, and include typically voice, signature and gait [1, 2]. Therefore, the choice of a biometric modality depends on several factors such as nonuniversality, nonpermanence, intraclass variations, poor image quality, noisy data, and matcher limitations [1, 3]. Thus, recognition based on unimodal biometric systems is not always reliable. To address these limitations, various works have been proposed for combining two or more biometric modalities in order to enhance the recognition performance [3, 4, 5]. This combination can be performed at data, feature, match score, and decision levels [3, 4].

However, with the existence of the constraints corresponding to the joint use of classifiers and methods of generating features, an appropriate operating method using mathematical approaches is needed, which takes into account two notions: uncertainty and imprecision of the classifier responses. In general, the most theoretical advances which have been devoted to the theory of probabilities are able to represent the uncertain knowledge but are unable to model easily the information that is imprecise, incomplete, or not totally reliable. Moreover, they often lead to confuse both concepts of uncertainty and imprecision with the probability measure. Therefore, new original theories dealing with uncertainty and imprecise information have been introduced, such as the fuzzy set theory [6], evidence theory [7], possibility theory [8] and, more recently, the theory of plausible and paradoxical reasoning developed by Dezert-Smarandache theory (DSmT) [9, 10, 11]. The DSmT has been elaborated by Jean Dezert and Florentin Smarandache for dealing with imprecise, uncertain and paradoxical sources of information. Thus, the main objective of the DSmT is to provide combination rules that would allow to correctly combine evidences issued from different information sources, even in presence of conflicts between sources or in presence of constraints corresponding to an appropriate model (i.e. free or hybrid DSm models [9]).

The use of the DSmT has been used justified in many kinds of applications [9, 10, 11]. Indeed, the DSmT is reported as very useful and powerful theoretical tool for enhancing the performance of multimodal biometric systems. Hence, combination algorithms based on DSmT have been used by Singh et al. [12] for robust face recognition through integrating multilevel image fusion and match score fusion of visible and infrared face images. Vatsa et al. proposed a DSmT based fusion algorithm [13] to efficiently combine level-2 and level-3 fingerprint features by incorporating image quality. Vatsa et al. proposed an unification of evidence-theoretic fusion algorithms [14] applied for fingerprint verification using level-2 and level-3 features. A DSmT based dynamic reconciliation scheme for fusion rule selection [15] has been proposed in order to manage the diversity of scenarios encountered in the probe dataset.

Generally, the handwritten signature is considered as the most known modality for biometric applications. Indeed, it is usually socially accepted for many government/legal/financial transactions such as validation of checks, historical documents, etc [16]. Hence, an intense research field has been devoted to develop various robust verification systems [17] according to the acquisition mode of the signature. Thus, two modes are used for capturing the signature, which are off-line mode and on-line mode, respectively. The off-line mode allows generating a handwriting static image from the scanning document. In contrast, the on-line mode allows generating dynamic information such as velocity and pressure from pen tablets or digitizers. For both modes, many Handwritten Signature Verification (HSV) systems have been developed in the past decades [17, 18, 19]. Generally, the off-line HSV systems remains less robust compared to the on-line HSV systems [16] because of the absence of dynamic information of the signer.

Generally, a HSV system is composed of three modules, which are preprocessing, feature generation and classification. In this context, various methods have been developed for improving the robustness of each individual HSV system. However, the handwritten signature verification failed to underline the incontestable superiority of a method over another in both steps of generating features and classification. Hence, rather than trying to optimize a single HSV system by choosing the best features for a given problem, researchers found more interesting to combine several classifiers [20].

Recently, approaches for combining classifiers have been proposed to improve signature verification performances, which led the development of several schemes in order to treat data in different ways [21]. Generally, three approaches for combining classifiers can be

considered: parallel approach [22, 23], sequential approach [24, 25] and hybrid approach [26], [27]. However, the parallel approach is considered as more simple and suitable since it allows exploiting the redundant and complementary nature of the responses issued from different signature verification systems. Hence, sets of classifiers have been used, which are based on global and local approaches [28, 29] and feature sets [30, 31], parameter features and function features [32, 33], static and dynamic features [34, 35]. Furthermore, several decision combination schemes have been implemented, ranging from majority voting [23, 36] to Borda count [37], from simple and weighted averaging [38] to Dempster-Shafer evidence theory [37, 39] and Neural Networks [40, 41]. The boosting algorithm has been used to train and integrate different classifiers, for both verification of on-line [42, 43] and off-line [44] signatures.

In this research, we follow the path of combined biometric systems by investigating the DSmT for combing different HSV systems. Therefore, we study the reliability of the DSmT for achieving a robust multiple HSV system. Two cases are considered for validating the effective use of the DSmT. The first one is to enhance the performance of off-line HSV systems by associating features based on Radon and Ridgelet transforms for each individual system. The second one is associating off-line image and dynamic information in order to improve the performance of single-source biometric systems and ensure greater security. For both cases, the combination is performed through the generalized biometric decision combination framework using Dezert-Smarandache theory (DSmT) [9, 10, 11].

The chapter is organized as follows. We give in Section 1.2 a review of sophisticated Proportional Conflict Redistribution (PCR5) rule based on DSmT. Section 1.3 describes the proposed verification system and Section 1.4 presents the performance criteria and datasets of handwritten signatures used for evaluation. Section 1.5 discuss the experimental results of the proposed verification system. The last section gives a summary of the proposed verification system and looks to the future research direction.

1.2 Review of PCR5 combination rule

Generally, the signature verification is formulated as a two-class problem where classes are associated to *genuine* and *impostor*, namely θ_{gen} and θ_{imp}, respectively. In the context of the probabilistic theory, the frame of discernment, namely Θ, is composed of two elements as: $\Theta = \{\theta_{gen}, \theta_{imp}\}$, and a mapping function $m \in [0,1]$ is associated for each class, which defines the corresponding mass verifying $m(\emptyset) = 0$ and $m(\theta_{gen}) + m(\theta_{imp}) = 1$.

When combining two sources of information and so two individual systems, namely information sources S^1 and S^2, respectively, the sum rule seems effective for non-conflicting responses [3]. In the opposite case, an alternative approach has been developed by Dezert and Smarandache to deal with (highly) conflicting imprecise and uncertain sources of information [9, 10, 11]. For two-class problem, a reference domain also called the frame of discernment should be defined for performing the combination, which is composed of a finite set of exhaustive and mutually exclusive hypotheses. Example of such approaches is PCR5 rule.

The main concept of the DSmT is to distribute unitary mass of certainty over all the composite propositions built from elements of Θ with $\cup$ (Union) and $\cap$ (Intersection) operators instead of making this distribution over the elementary hypothesis only. Therefore, the hyper-powerset D^Θ is defined as $D^\Theta = \{\emptyset, \theta_{gen}, \theta_{imp}, \theta_{gen} \cup \theta_{imp}, \theta_{gen} \cap \theta_{imp}\}$. The DSmT uses the generalized basic belief mass, also known as the generalized basic belief assignment (gbba) computed on hyper-powerset of Θ and defined by a map $m(.): D^\Theta \longrightarrow [0,1]$ associated to a given source of evidence, which can support paradoxical information, as follows: $m(\emptyset) = 0$ and $m(\theta_{gen}) + m(\theta_{imp}) + m(\theta_{gen} \cup \theta_{imp}) + m(\theta_{gen} \cap \theta_{imp}) = 1$. The

combined masses m_{PCR5} obtained from $m_1(.)$ and $m_2(.)$ by means of the PCR5 rule [10] is defined as:

$$m_{PCR5}(A) = \begin{matrix} 0 & \text{if } A \in \Phi \\ m_{DSmC}(A) + m_{A\cap X}(A) & \text{otherwise} \end{matrix} \quad (1.1)$$

Where

$$m_{A\cap X}(A) = \underset{\substack{X\in D^\Theta\setminus\{A\} \\ c(A\cap X)=\emptyset}}{} \frac{m_1(A)\}^2\, m_2(X}{m_1(A)+m_2(X)} + \frac{m_2(A)\}^2\, m_1(X}{m_2(A)+m_1(X)}$$

and $\Phi = \{\Phi_{\mathcal{M}}, \emptyset\}$ is the set of all relatively and absolutely empty elements, $\Phi_{\mathcal{M}}$ is the set of all elements of D^Θ which have been forced to be empty in the Shafer's model $\mathcal{M}$ defined by the exhaustive and exclusive constraints, $\emptyset$ is the empty set, and $c(A \cap X)$ is the canonical form (conjunctive normal) of $A \cap X$ and where all denominators are different to zero. If a denominator is zero, that fraction is discarded. Thus, the term $m_{DSmC}(A)$ represents a conjunctive consensus, also called DSm Classic (DSmC) combination rule [9], which is defined as:

$$m_{DSmC}(A) = \begin{matrix} 0 & \text{if } A = \emptyset \\ \underset{(X,Y\in D^\Theta, X\cap Y=A)}{} m_1(X)m_2(X) & \text{otherwise} \end{matrix} \quad (1.2)$$

1.3 System description

The combined individual HSV system is depicted in Figure 1.1, which are composed of an off-line verification system, an on-line or off-line verification system and a combination module. s_1 and s_2 define the off-line and on-line or off-line handwritten signatures provided by two sources of information S^1 and S^2, respectively. Each individual verification system is generally composed of three modules: pre-processing, feature generation and classification.

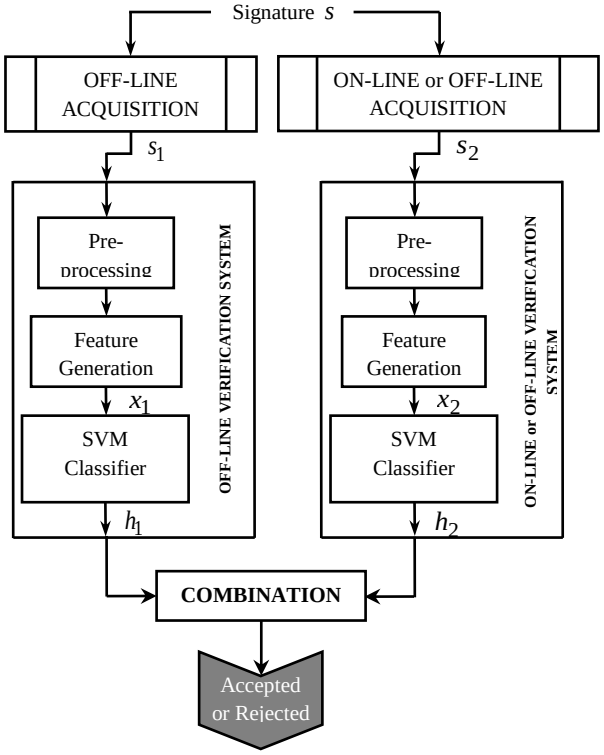


Figure 1.1: Structure of the combined individual HSV systems.

1.3.1 Pre-processing

According the acquisition mode, each handwritten signature is pre-processed for facilitating the feature generation. Hence, the pre-processing of the off-line signature includes two steps: Binarization using the local iterative method [45] and elimination of the useless information around the signature image without unifying its size. The pre-processing steps of a signature example are shown in Figure 1.2. The binarization method was chosen to capture signature from the background. It takes the advantages of locally adaptive binarization methods [45] and adapts them to produce an algorithm that thresholds signatures in a more controlled manner. By doing this, the local iterative method limits the amount of noise generated, as well as it attempts to reconstruct sections of the signature that are disjointed.

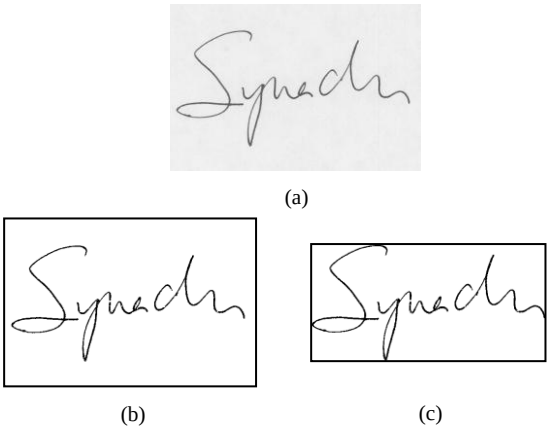


Figure 1.2: Preprocessing steps: (a) Scanning (b) Binarization
(c) Elimination of the useless information.

While the on-line signature, no specific pre-processing is required. More details on the acquisition method and pre-processing module of the on-line signatures are provided in Refs. [46] and [47].

1.3.2 Feature generation

Features are generated according the acquisition mode. In the combined individual HSV systems, we use the uniform grid, Radon and Ridgelet transforms for off-line signatures and dynamic characteristics for on-line signatures, respectively.

a. Features used for combining individual off-line HSV systems

The first case study for evaluating the performance of the proposed combination using DSmT is performed with two individual off-line HSV systems. Features are generated from the same off-line signature using the Radon and Ridgelet transforms. The Radon transform is well adapted for detecting linear features. In contrast, the Ridgelet transform allows representing linear singularities [48]. Therefore, Radon and Ridgelet coefficients provide complementary information about the signature.

- **Radon transform based features:** The Radon transform of each off-line signature is calculated by setting the respective number of projection points N_r and orientations N_θ, which define the length of the radial and angular vectors, respectively. Hence, a radon matrix is obtained having a size $[N_r \times N_\theta]$ which provides in each point the cumulative

intensity of pixels forming the image of the off-line signature. Figure 1.3 shows an example of a binarized image of an off-line signature and the steps involved for generating features based on Radon transform. Since the Radon transform is redundant, we take into account only positive radial points $\}\}N_r/2\}\times N_\theta\}$. Then after, for each angular direction, the energy of Radon coefficients is computed to form the feature vector x_1 of dimension $\}1\times N_\theta\}$. This energy is defined as:

$$E_\theta^{rad} = \frac{2}{N_r}\sum_{r=1}^{N_r/2} T_{rad}^2(r,\theta \quad ,\theta \in \{1,2,\dots,N_\theta\} \tag{1.3}$$

where T_{rad} is the Radon transform operator.

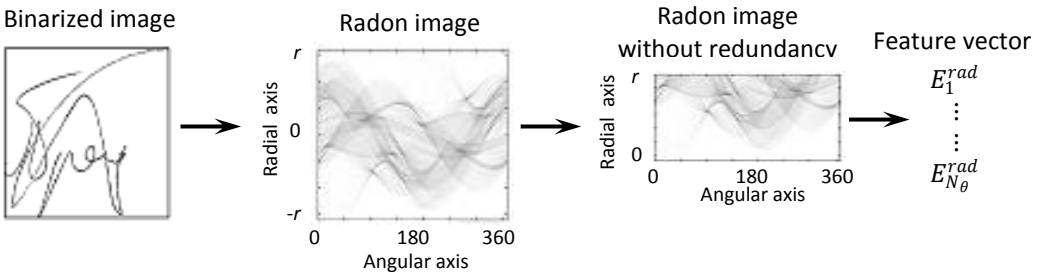


Figure 1.3: Steps for generating the feature vector from the Radon transform.

- **Ridgelet transform based features:** For generating complementary information of the Radon features, the wavelet transform (WT) is performed along the radial axis allowing generating the Ridgelet coefficients [49]. Figure 1.4 shows an example for generating the feature vector from the Ridgelet transform. For each angular direction, the energy of Ridgelet coefficients is computed taking into account only details issued from the decomposition level L of the WT. Therefore, the different values of energy are finally stored in a vector x_2 of dimension $\}1\times N_\theta\}$. This energy is defined as:

$$\frac{2}{N_r}\sum_{r=1}^{N_r/2} E_\theta^{rid} = \}{T_r}^2{}_{id}\}a,b,\theta\},\theta\in\}1,2,\dots,N_\theta\} \tag{1.4}$$

where T_{rid} is the Ridgelet transform operator whereas a and b are the scaling and translation factors of the WT, respectively.

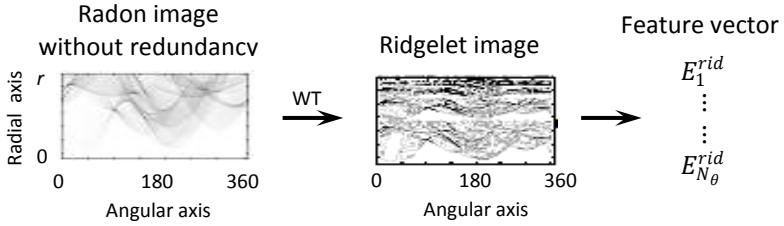


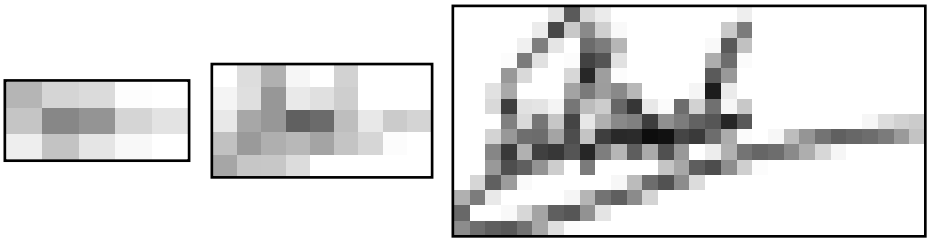
Figure 1.4: Steps for generating the feature vector from the Ridgelet transform.

b. Features used for combining individual off-line and on-line HSV systems

The second case study is considering for evaluating the performance of the proposed DSmT for combining both individual off-line and on-line HSV systems. Features are generated from both off-line and on-line signatures of the same user using the uniform grid (UG) and dynamic characteristics, respectively. The UG allows extracting locally features without normalization of the off-line signature image. On each grid, the densities are

computed providing overall signature appearance information. In contrast, dynamic characteristics computed from the on-line signature allow providing complementary dynamic information in the combination process.

- **Uniform grid based features:** Features are generated using the Uniform Grid (UG) [50, 51], which consists to create $n \times m$ rectangular regions for sampling. Each region has the same size and shape. Parameters n and m define the number of lines (vertical regions) and columns (horizontal regions) of the grid, respectively. Hence, the feature associated to each region is defined as the ratio of the number of pixels belonging to the signature and the total number of pixels of images. Therefore, the different values are finally stored in a vector x_1 of dimension $n \times m$, which characterizes the off-line signature image.
  Figure 1.5 shows a 3×5 grid, which allows an important reduction of the representation vector, but it preserves wrongly the visual information. In contrast, a 15×30 grid which provides an accurate representation of images, but it leads a larger characteristic vector. A 5×9 grid seems to be an optimal choice between the quality of representation and dimensionality. Thus, the optimal choice of the grid size for all writers is obviously too important to effectively solve our problem of signature verification. In our case, for all experiments, the parameters n and m of are fixed to 5 and 9, respectively.



$[3\times5]$ $[5\times9]$ $[15\times30]$

Figure 1.5: Visualization of different grid sizes.

- **Dynamic information based features:** For the individual on-line verification system, features are generated using only the dynamic features. Each on-line signature is represented by a vector x_2 composed of 11 features, which are signature total duration, average velocity, vertical average velocity, horizontal average velocity, maximal velocity, average acceleration, maximal acceleration, variance of pressure, mean of azimuth angle, variance of azimuth angle and mean of elevation angle. A complete description of the feature set is shown in Table 1.1.

1.3.3 Classification based on SVM

a. Review of SVMs

The classification based on Support Vector Machines (SVMs) has been widely used in many pattern recognition applications as the handwritten signature verification [35, 52]. The SVM is a learning method introduced by Vapnik et al. [53], which tries to find an optimal hyperplane for separating two classes. Its concept is based on the maximization of the distance of two points belonging each one to a class. Therefore, the misclassification error of data both in the training set and test set is minimized.

Ranking	Feature Description	Ranking	Feature Description
1	$t_n - t_1$	7	$\max_{i=1,\ldots,n-2} \frac{dist_{Eucl}(Pt_i, Pt_{i+2})}{(t_{i+1} - t_i)^2}$
2	$\frac{\sum_{i=1}^{n-1}(dist_{Eucl}(Pt_i, Pt_{i+1})}{t_n - t_1}$	8	$\sum_{i=1}^{n}\left(Pr_i - \frac{\sum_{i=1}^{n} Pr_i}{n}\right)^2$
3	$\frac{\sum_{i=1}^{n-1}\lvert y_{i+1} - y_i}{t_n - t_1}$	9	$\frac{\sum_{i=1}^{n} Az_i}{n}$
4	$\frac{\sum_{i=1}^{n-1}\lvert x_{i+1} - x_i}{t_n - t_1}$	10	$\sum_{i=1}^{n}\left(Az_i - \frac{\sum_{i=1}^{n} Az_i}{n}\right)^2$
5	$\max_{i=1,\ldots,n-1} \frac{dist_{Eucl}(Pt_i, Pt_{i+1})}{t_{i+1} - t_i}$	11	$\frac{\sum_{i=1}^{n} Al_i}{n}$
6	$\frac{\sum_{i=1}^{n-1}(dist_{Eucl}(Pt_i, Pt_{i+1}}{(t_n - t_1)^2}$		

Table 1.1: Set of dynamic features. $s =)Pt_1, Pt_2, \ldots, Pt_n)$ denotes an on-line signature composed of n events $Pt_i)x_i, y_i, t_i)$, $x_i, y_i, Pr_i, Az_i, Al_i$ denote x-position, y-position, pen pressure, azimuth and elevation angles of the pen at the i^{th} time instant t_i, respectively.

Basically, SVMs have been defined for separating linearly two classes. When data are non linearly separable, a kernel function is used. Thus, all mathematical functions, which satisfy Mercer's conditions, are eligible to be a SVM-kernel [53]. Examples of such kernels are sigmoid kernel, polynomial kernel, and Radial Basis Function (RBF) kernel. Generally, the RBF kernel is used for its better performance, which is defined as:

$$K(x, x_k) = exp\left(-\frac{x - x_k\|^2}{2\,\sigma^2}\right. \tag{1.5}$$

Where σ is the kernel parameter, $\|x - x_k\|$ is the Euclidian distance between two samples. Therefore, the decision function f: $\mathbb{R}^p \rightarrow \{-1, +1\}$, is expressed in terms of kernel expansion as:

$$f(x) = \sum_{k=1}^{Sv} \alpha_k y_k\, K(x, x_k) + b \tag{1.6}$$

where α_k are Lagrange multipliers, Sv is the number of support vectors x_k which are training data, such that $0 \leq \alpha_k \leq C$, C is a user-defined parameter that controls the tradeoff between the machine complexity and the number of nonseparable points [54], the bias b is a scalar computed by using any support vector. Finally, test data x_d, $d = \{1,2\}$, are classified according to:

$$x_d \in \begin{matrix} class\ (+1) & \text{if } f(x_d) > 0 \\ class\ (-1) & \text{otherwise} \end{matrix} \tag{1.7}$$

b. Decision rule

The direct use of SVMs does not allow defining a decision threshold to assign a signature to genuine or forgery classes. Therefore, outputs of SVM are transformed to objective evidences, which express the membership degree (MD) of a signature to both classes (genuine or forgery). In practice, the MD has no standard form. However, the only constraint is that it must be limited in the range of $[0, 1]$ whereas SVMs produce a single output. In this chapter, we use a fuzzy model which has been proposed in [50, 51, 55] to assign MD for SVM output in both genuine and impostor classes. Let $f(x_d)$ be the output of a SVM obtained for a given signature to be classified. The respective membership degrees $h_d(\theta_i), i = \{gen, imp\}$

associated to genuine and impostor classes are defined according to membership models given in the Algorithm 1 [51]. To compute the values of membership degrees $)h_d, d = 1, 2)$, we consider the two case studies as follows:

- in first case study, the main problem for generating features is the appropriate number of the angular direction N_θ for the Radon transform and the number of the decomposition level L of the WT (Haar Wavelet) in the Ridgelet domain. Hence, many experiments are conducted for finding the optimal values for which the error rate in the training phase is null. In this case, feature vectors are generated from both Radon $(d = 1)$ and Ridgelet $(d = 2)$ of the same off-line signature by setting N_θ and L to 32 and 3, respectively.
- in second case study, we calculate the values $(h_d, d = 1)$ of off-line signature by using the optimal size $[5 \times 9]$ of the grid for which the error rate in the training phase is null. In the same way, we calculate also the values $(h_d, d = 2)$ of on-line signature by using the vector of 11 dynamic features for which the error rate in the training phase is null.

Algorithm 1.

Respective membership models for two classes.

if $f(x_d) > 1$ then
  $h_d(\theta_{gen}) \leftarrow 1$
  $h_d(\theta_{imp}) \leftarrow 0$
else
  if $f(x_d) < -1$ then
    $h_d(\theta_{gen}) \leftarrow 0$
    $h_d(\theta_{imp}) \leftarrow 1$
  else
    $h_d(\theta_{gen}) \leftarrow \frac{1 + f(x_d)}{2}$
    $h_d(\theta_{imp}) \leftarrow \frac{1 - f(x_d)}{2}$
  end if
end if

Hence, a decision rule is performed about whether the signature is genuine or forgery as described in Algorithm 2.

Algorithm 2. Decision making in SVM framework.

if $\frac{h_d(\theta_{gen})}{h_d(\theta_{imp})} \geq t$ then
    $s_d \in \theta_{gen}$
else
    $s_d \in \theta_{imp}$
end if

Where t defines a decision threshold.

1.3.4 Classification based on DSmT

The proposed combination module consists of three steps: i) transform membership degrees of the SVM outputs into belief assignments using estimation technique based on the dissonant model of Appriou, ii) combine masses through a DSmT based combination rule and iii) make a decision for accepting or rejecting a signature.

a. Estimation of masses

In this chapter, the mass functions are estimated using a dissonant model of Appriou, which is defined for two classes [56]. Therefore, the extended version of Appriou's model in DSmT framework is given as:

$$m_{id}(\emptyset) = 0 \tag{1.8}$$

$$m_{id}(\theta_i) = \frac{(1-\beta_{id})\, h_d(\theta_i)}{1+h_d(\theta_i)} \tag{1.9}$$

$$m_{id}(\overline{\theta_i}) = \frac{1-\beta_{id}}{1+h_d(\theta_i)} \tag{1.10}$$

$$m_{id}(\theta_i \cup \overline{\theta_i}) = \beta_{id} \tag{1.11}$$

$$m_{id}(\theta_i \cap \overline{\theta_i}) = 0 \tag{1.12}$$

where $i = \{gen, imp\}$, $h_d(\theta_i)$ is the membership degree of a signature provided by the corresponding source S^d $(d = 1, 2)$, $(1 - \beta_{id})$ is a confidence factor of i-th class, and β_{id} defines the error provided by each source $(d = 1, 2)$ for each class θ_i. In our approach, we consider β_{id} as the verification accuracy prior computed on the training database for each class [14]. Since both SVM models have been validated on the basis that errors during training phase are zero, therefore β_{id} is fixed to 0.001 in the estimation model.

Note that the same information source cannot provide two responses, simultaneously. Hence, in DSmT framework, we consider that the paradoxical hypothesis $\theta_i \cap \theta\}_i$ has no physical sense towards the two information sources θ_{gen} and θ_{imp} . Therefore, the beliefs assigned to this hypothesis are null as given in Equation (1.12).

b. Combination of masses

The combined masses are computed in two steps. First, the belief assignments $(m_{id}(.), i = \{gen, imp\})$ are combined for generating the belief assignments for each source as follows:

$$m_1 = m_{\{gen\}1} \oplus m_{\{imp\}1} \tag{1.13}$$

$$m_2 = m_{\{gen\}2} \oplus m_{\{imp\}2} \tag{1.14}$$

where $\oplus$ represents the conjunctive consensus of the DSmC rule.

Finally, the belief assignments for the combined sources $(m_d(.), d = 1, 2)$ are then computed as:

$$m_c = m_1 \oplus m_2 \tag{1.15}$$

where $\oplus$ represents the combination operator, which is composed of both conjunctive and redistribution terms of the PCR5 rule.

c. Decision rule

A decision for accepting or rejecting a signature is made using the statistical classification technique. First, the combined beliefs are converted into probability measure using a probabilistic transformation, called Dezert-Smarandache probability (DSmP), that maps a belief measure to a subjective probability measure [11] defined as:

$$DSmP_{\epsilon}(\theta_i) = m_c(\theta_i) + (m_c(\theta_i) + \epsilon)w_{\mathcal{M}} \tag{1.16}$$

where $w_{\mathcal{M}}$ is a weighting factor defined as:

$$w_{\mathcal{M}} = \underset{\substack{A_j \in 2^{\Theta} \\ A_j \supset \theta_i \\ C_{\mathcal{M}}(A_j) \geq 2}}{} \frac{m_c(A_j)}{\underset{\substack{A_k \in 2^{\Theta} \\ A_k \subset X \\ C_{\mathcal{M}}(A_k)=1}}{} m_c(A_k) + \epsilon\, C_{\mathcal{M}}(A_j}$$

such that is a tuning parameter, $\mathcal{M}$ is the Shafer's model for Θ, and $C_{\mathcal{M}})A_k)$ denotes the DSm cardinal [11] of the set A_k . Therefore, the likelihood ratio test is performed for decision making as described in Algorithm 3.

Algorithm 3. Decision making in DSmT framework.

if $\frac{DSmP_{\varepsilon}(\theta_{gen})}{DSmP_{\varepsilon}(\theta_{imp})} \geq t$ then
    $s_d \in \theta_{gen}$
else
    $s_d \in \theta_{imp}$
end if

Where t defines a decision threshold and $s = \{s_1, s_2\}$ is the j -th signature represented by two modalities according the case study as follows:

- in first case study, s is an off-line signature characterized by both Radon and Ridgelet features.
- in second case study, s is a signature represented by both off-line and on-line modalities.

1.4 Performance criteria and dataset description

In this section, we briefly describe datasets used and performance criteria for evaluating the proposed DSmT for combing individual handwritten signature verification systems.

1.4.1 Dataset description

To evaluate the verification performance of the proposed DSmT based combination of individual HSV systems, we use two datasets of handwritten signatures: (1) CEDAR signature dataset [57] used for evaluating the performance for combining individual off-line HSV systems and (2) NISDCC signature dataset [58] for the experiments related to the simultaneous verification of individual off-line and on-line HSV systems.

a. CEDAR signature database

The Center of Excellence for Document Analysis and Recognition (CEDAR) signature dataset [57] is a commonly used for off-line signature verification. The CEDAR dataset consists of 55 signature sets, each one being composed by one writer. Each writer provided 24 samples of their signature, where these samples constitute the genuine portion of the dataset. While, forgeries are obtained by asking arbitrary people to skillfully forge the signatures of the previously mentioned writers. In this fashion, 24 forgery samples are collected per writer from about 20 skillful forgers. In total, this dataset contains 2640 signatures, built from 1320 genuine signatures and 1320 skilled forgeries. Figures 1.6(a) and 1.6(b) show two examples of both preprocessed genuine and forgery signatures for one writer, respectively.

(a) Genuine signatures. (b) Forgery signatures.

Figure 1.6: Signature samples of the CEDAR.

b. NISDCC signature database

The Norwegian Information Security laboratory and Donders Centre for Cognition (NISDCC) signature dataset has been used in the ICDAR'09 signature verification competition [58]. This collection contains simultaneously acquired on-line and off-line samples. The off-line dataset is called "NISDCC-offline" and contains only static information while the on-line dataset which is called "NISDCC-online" also contains dynamic information, which refers to the recorded temporal movement of handwriting process. Thus, the acquired on-line signature is available under form of a subsequent sampled trajectory points. Each point is acquired at 200 Hz on tablet and contains five recorded pen-tip coordinates: x-position, y-position, pen pressure, azimuth and elevation angles of the pen [59] The NISDCC-offline dataset is composed of 1920 images from 12 authentic writers (5 authentic signatures per writer) and 31 forging writers (5 forgeries per authentic signature). Figures 1.7(a) and 1.7(b) show an example of both preprocessed off-line signature and a plotted matching on-line signature for one writer, respectively.

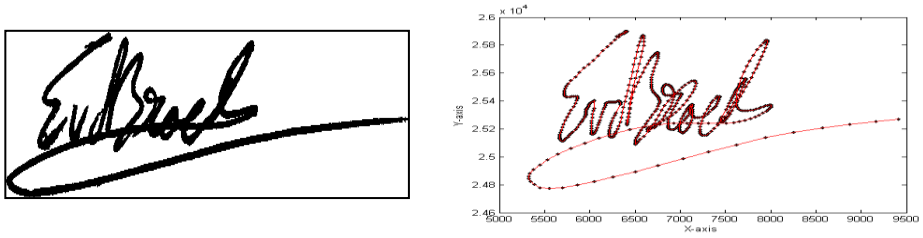


Figure 1.7: Signature samples of the NISDCC signature collection.

1.4.1 Performance criteria

For evaluating performances of the combined individual HSV systems, three different kinds of error are considered: False Accepted Rate (FAR) allows taking into account only skilled forgeries; False Rejected Rate (FRR) allows taking into account only genuine signatures and finally the Half Total Error Rate (HTER) allows taking into account both rates. Thus, Equal Error Rate is a special case of HTER when FRR = FAR.

1.4.2 SVM model

For both case studies, signature data are split into training and testing sets for evaluating the performances of the proposed DSmT based combination of individual HSV systems. Thus, the training phase allows finding the optimal hyperparameters for each individual SVM model. In our system, the RBF kernel is selected for the experiments.

a. SVM models used for combined individual off-line HSV systems

In first case study, the SVM model is produced for each individual off-line HSV system according the Radon and Ridgelet features, respectively. For each writer, 2/3 and 1/3 samples are used for training and testing, respectively. The optimal parameters (C, σ) of each SVM are tuned experimentally, which are fixed as $(C = 19.1, \sigma = 4)$ and $(C = 15.1, \sigma = 4.6)$, respectively.

b. SVM models used for combined individual off-line and on-line HSV systems

In second case study, the SVM model is produced for both individual off-line and on-line HSV systems according the uniform grid features and dynamic information, respectively. For each writer and both datasets, 2/3 and 1/3 samples are used for training and testing, respectively. The optimal parameters (C, σ) for both SVM classifiers (off-line and on-line) are tuned experimentally, which are fixed as $(C = 9.1, \sigma = 9.4)$ and $(C = 13.1, \sigma = 2.2)$, respectively.

1.5 Experimental results and discussion

For each case study, decision making will be only done on the simple classes. Hence, we consider the masses associated to all classes belonging to the hyper power set $D^{\Theta} = \{\emptyset, \theta_{gen}, \theta_{imp}, \theta_{gen} \cup \theta_{imp}, \theta_{gen} \cap \theta_{imp}\}$ in both combination process and decision making. In the context of signature verification, we take as constraint the proposition that $\theta_{gen} \cap \theta_{imp} = \emptyset$ in order to separate between genuine and impostor classes. Therefore, the hyper power set D^{Θ} is simplified to the power set 2^{Θ} as $2^{\Theta} = \{\emptyset, \theta_{gen}, \theta_{imp}, \theta_{gen} \cup \theta_{imp}\}$, which defines the Shafer's model [9]. This section presents the experimental results with their discussion.

To evaluate the performance of the proposed DSmT based combination, we use two individual off-line HSV systems using the CEDAR database at the first case study. Indeed, the task of the proposed combination module is to manage the conflicts generated between the two individual off-line HSV systems for each signature using the PCR5 combination rule. For that, we compute the verification errors of both individual off-line HSV systems and the combined individual off-line HSV systems using PCR5 rule. Figure 1.8 shows the FRR and FAR computed for different values of decision threshold using both individual off-line HSV systems of this first case study. Table 1.2 shows the verification errors rates computed for the corresponding optimal values of decision threshold of this case study. Here HSV system 1 is the individual off-line verification system feeded by Radon features that yields an error rate of 7.72% corresponding to the optimal value of threshold $t = 1.11$ while HSV system 2 is the individual off-line verification system feeded by Ridgelet features, which provides the same result with an optimal value of threshold $t = 0.991$. Consequently, both individual off-line HSV systems give the same verification performance since the corresponding error rate of $\text{HTER} = 7.72\%$ is the same.

The proposed DSmT based combination of individual off-line HSV systems yields a HTER of 5.45% corresponding to the optimal threshold value $t = 0.986$. Hence, the combined individual off-line HSV systems with PCR5 rule allows improving the verification performance by 2.27%. This is due to the efficient redistribution of the partial conflicting mass only to the elements involved in the partial conflict.

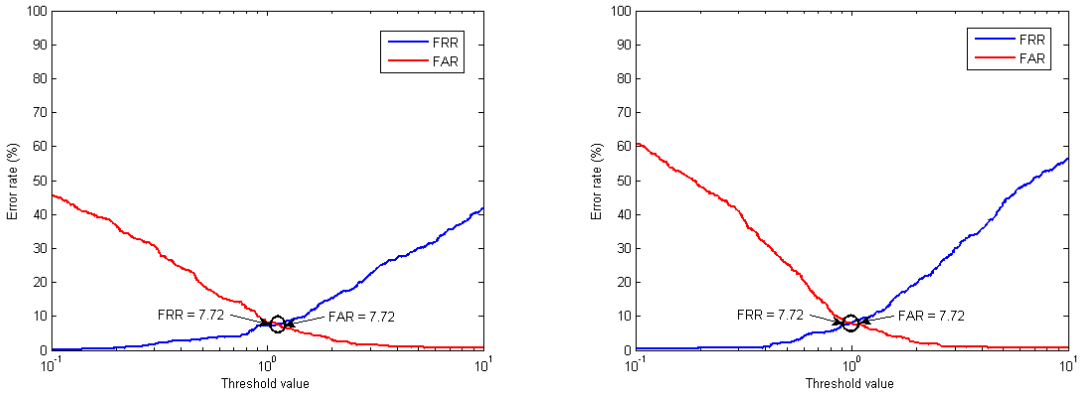


Figure 1.8: Performance evaluation of the individual off-line HSV systems.

HSV Systems	Optimal Threshold	FAR	FRR	HTER
System 1	1.110	7.72	7.72	7.72
System 2	0.991	7.72	7.72	7.72
Combined Systems	0.986	5.45	5.45	5.45

Table 1.2: Error rates (%) obtained for individual and combined HSV systems.

In the second case study, two sources of information are combined through the PCR5 rule. Figure 1.9 shows three examples of conflict measured between off-line and on-line signatures for writers 3, 7, and 10 of the NISDCC dataset, respectively. The values K_{c3})∈)0.00, 0.35)), K_{c7})∈)0.00, 0.64)), and K_{c10})∈)0.00, 1.00)) represent the mass assigned to the empty set, after combination. We can see that the two sources of information are very conflicting. Hence, the task of the proposed combination module is to manage the conflicts generated from both sources $)K_{cw}, w = 1, 2, \dots, 12)$ for each signature using the PCR5 combination rule. For that, we compute the verification errors of both individual off-line and on-line HSV systems and the proposed DSmT based combination. Figure 1.10 shows the FRR and FAR computed for different values of decision threshold using both individual off-line and on-line HSV systems of this second case study. For better comparison, Table 1.3 shows the HTER computed for the corresponding optimal values of decision threshold of this case study.

The proposed DSmT based combination of both individual off-line and on-line HSV systems yields a HTER of 0% corresponding to the optimal threshold value $t = 0.597$. Consequently, the proposed combination of individual off-line and on-line HSV systems using PCR5 rule yields the best verification accuracy compared to the individual off-line and on-line HSV systems, which provide conflicting and complementary outputs.

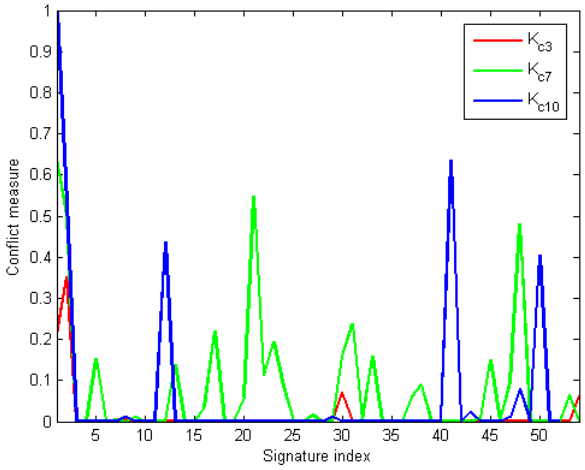
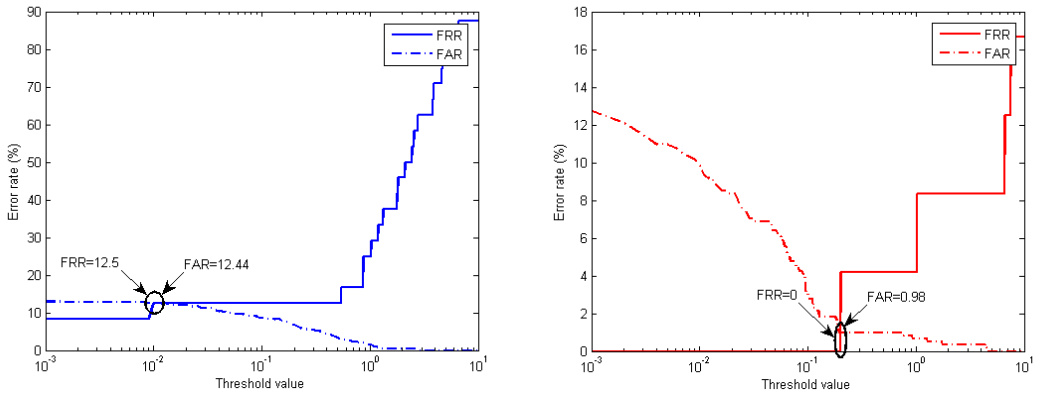


Figure 1.9: Conflict between off-line and on-line signatures
for the writers 3, 7, and 10, respectively.



(a) Off-line HSV system 1.

(b) On-line HSV system 2.

Figure 1.10: Performance evaluation of the individual off-line and on-line HSV systems.

HSV Systems	Optimal Threshold	FAR	FRR	HTER
System 1	0.012	12.44	12.50	12.47
System 2	0.195	0.98	0.00	0.49
Combined Systems	0.597	0.00	0.00	0.00

Table 1.3: Error rates (%) obtained for individual and combined HSV systems.

1.6 Conclusion

This chapter proposed and presented a new system based on DSmT for combining different individual HSV systems which provide conflicting results. The individual HSV systems are combined through DSmT using the estimation technique based on the dissonant model of Appriou, sophisticated PCR5 rule and likelihood ratio test. Hence, two cases have been addressed in order to ensure a greater security: (1) combining two individual off-line HSV systems by associating Radon and Ridgelet features of the same off-line signature (2) and combining both individual off-line and on-line HSV systems by associating static image

and dynamic information of the same signature characterized by off-line and on-line modalities. Experimental results show in both case studies that the proposed system using PCR5 rule allows improving the verification errors compared to the individual HSV systems.

As remark, although the DSmT allows improving the verification accuracy in both studied cases, it is clearly that the achieved improvement depends also to the complementary outputs provided by the individual HSV systems. Indeed, according to the second case study, a suitable performance quality on the individual on-line HSV system can be obtained when the dynamic features of on-line signatures are carefully chosen. Combined to the grid features using DSmT allows providing more powerful system comparatively to the system of the first case study in term of success ratio. In continuation to the present work, the next objectives consist to explore other alternative DSmT based combinations of HSV systems in order to attempt improving performance quality of the writer-independent HSV whether the signature is genuine or forgery as well as in the false rejection and false acceptance concepts.

1.7 References

[1] A.K. Jain, P. Flynn and A. Ross, *Handbook of Biometrics*, Springer-Verlag, New York, 2007.

[2] A.K. Jain, A. Ross and S. Prabhakar, *An introduction to biometric recognition*, IEEE Transaction on Circuits and Systems for Video Technology, Special Issue on Image- and Video- Based Biometrics, Vol. 14(1), pp. 4–20, 2004.

[3] A. Ross, K. Nandakumar and A. K. Jain, *Handbook of Multibiometrics*, Springer-Verlag, New York, 2006.

[4] A. Ross and A.K. Jain, *Information fusion in biometrics*, Pattern Recognition Letters, Vol. 24(13), pp. 2115–2125, 2003.

[5] J. Kittler, M. Hatef, R.P. Duin and J.G. Matas, *On combining classifiers*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20(3), pp. 226–239, 1998.

[6] L.A. Zadeh, *Fuzzy algorithm*, Information and Control, Vol. 12, pp. 94-102, 1968.

[7] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[8] D. Dubois and H. Prade, *Representation and combination of uncertainty with belief functions and possibility measures*, Computational Intelligence, Vol. 4, pp. 244-264, 1988.

[9] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion*, Rehoboth, NM: Amer. Res. Press, 2004.

[10] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion*, Rehoboth, NM: Amer. Res. Press, 2006.

[11] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion*, Rehoboth, NM: Amer. Res. Press, 2009.

[12] R. Singh, M. Vatsa and A. Noore, *Integrated Multilevel Image Fusion and Match Score Fusion of Visible and Infrared Face Images for Robust Face Recognition*, Pattern Recognition - Special Issue on Multimodal Biometrics, Vol. 41(3), pp. 880-893, 2008.

[13] M. Vatsa, R. Singh, A. Noore, and M. Houck, *Quality-Augmented Fusion of Level-2 and Level-3 Fingerprint Information using DSm Theory*, International Journal of Approximate Reasoning, Vol. 50(1), 2009.

[14] M. Vatsa, R. Singh and A. Noore, *Unification of Evidence Theoretic Fusion Algorithms: A Case Study in Level-2 and Level-3 Fingerprint Features*, IEEE Transaction on Systems, Man, and Cybernetics - A, Vol 29(1), 2009.

[15] M. Vatsa, R. Singh, A. Ross and A. Noore, *On the Dynamic Selection in Biometric Fusion Algorithms*, IEEE Transaction on Information Forensics and Security, Vol. 5(3), pp. 470-479, 2010.

[16] D. Impedovo and G. Pirlo, *Automatic Signature Verification: The State of the Art*, IEEE Transactions on Systems, Man, and Cybernetics-C, 38(5), pp. 609–335, 2008.

[17] R. Plamondon and S.N. Srihari, *On-line and off-line handwriting recognition: A comprehensive survey*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22(1), pp. 63–84, 2000.

[18] R. Plamondon and G. Lorette, *Automatic signature verification and writer identification: The state of the art*, Pattern Recognition, Vol. 22(2), pp. 107-131, 1989.

[19] F. Leclerc and R. Plamondon, *Automatic signature verification: The state of the art 1989-1993*, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 8(3), pp. 643-660, 1994.

[20] D. Ruta and B. Gabrys, *An overview of classifier fusion methods*, Computing and Information Systems, Vol. 7(1), pp. 1-10, 1994.

[21] L.P. Cordella, P. Foggia, C. Sansone, F. Tortorella, and M. Vento, *Reliability parameters to improve combination strategies in multi-expert systems*, Pattern Analysis and Application, Vol. 3(2), pp. 205–214, 1999.

[22] Y. Qi and B.R. Hunt, *A multiresolution approach to computer verification of handwritten signatures*, IEEE Transactions on Image Processing, Vol. 4(6), pp. 870–874, 1995.

[23] G. Dimauro, S. Impedovo, G. Pirlo and A. Salzo, *A multi-expert signature verification system for bankcheck processing*, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 11(5), pp. 827–844, 1997, [*Automatic Bankcheck Processing* (Series in Machine Perception and Artificial Intelligence), Vol. 28, S. Impedovo, P. S. P. Wang and H. Bunke, Eds. Singapore: World Scientific, pp. 365–382].

[24] C. Sansone and M. Vento, *Signature verification: Increasing performance by a multi-stage system*, Pattern Analysis and Application, Vol. 3, pp. 169-181, 2000.

[25] K. Zhang, E. Nyssen and H. Sahli, *A multi-stage online signature verification system*, Pattern Analysis and Application, Vol. 5, pp. 288-295, 2002.

[26] L.P. Cordella, P. Foggia, C. Sansone and M. Vento, *Document validation by signature: A serial multi-expert approach*, in Proceedings of 5th International Conference on Document Analysis and Recognition, pp. 601–604, 1999.

[27] L.P. Cordella, P. Foggia, C. Sansone, F. Tortorella and M. Vento, *A cascaded multiple expert system for verification*, in Proceedings of 1st International Workshop, Multiple Classifier Systems, (Lecture Notes in Computer Science), Vol. 1857, J. Kittler and F. Roli, Eds. Berlin, Germany: Springer-Verlag, pp. 330–339, 2000.

[28] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Penalba, J. Ortega-Garcia and D. Maltoni, *An on-line signature verification system based on fusion of local and global information*, (Lecture Notes in Computer Science 3546), in Audio- and Video-Based Biometric Person Authentication, New York: Springer-Verlag, pp. 523–532, 2005.

[29] S. Kumar, K.B. Raja, R.K. Chhotaray and S. Pattanaik, *Off-line Signature Verification Based on Fusion of Grid and Global Features Using Neural Networks*, International Journal of Engineering Science and Technology, Vol. 2(12), pp. 7035-7044, 2010.

[30] K. Huang and H. Yan, *Identifying and verifying handwritten signature images utilizing neural networks*, in Proceedings ICONIP, pp. 1400–1404, 1996.

[31] K. Huang, J. Wu and H. Yan, *Offline writer verification utilizing multiple neural networks*, Optical Engineering, Vol. 36(11), pp. 3127–3133, 1997.

[31] R. Plamondon, P. Yergeau and J.J. Brault, *A multi-level signature verification system*, in From Pixels to Features III—Frontiers in Handwriting Recognition, S. Impedovo and J. C. Simon, Eds. Amsterdam, The Netherlands: Elsevier, pp. 363–370, 1992.

[32] I. Nakanishi, H. Hara, H. Sakamoto, Y. Itoh and Y. Fukui, *Parameter Fusion in DWT Domain: On-Line Signature Verification*, in International Symposium in Intelligent Signal Processing and Communication Systems, Yonago Convention Center, Tottori, Japan, 2006.

[33] M. Liwicki, Y. Akira, S. Uchida, M. Iwamura, S. Omachi and K. Kise, *Reliable Online Stroke Recovery from Offline Data with the Data-Embedding Pen*, in Proceedings of 11th International Conference Document Analysis and Recognition, pp. 1384-1388, 2011.

[34] V. Mottl, M. Lange V. Sulimova and A. Yermakov, *Signature verification based on fusion of on-line and off-line kernels*, in Proceedings of 19-th International Conference on Pattern Recognition, Florida, USA, December 08-11, 2008.

[35] V.E. Ramesh and M.N. Murty, *Offline signature verification using genetically optimized weighted features*, Pattern Recognition, Vol. 32(2), pp. 217–233, 1999.

[36] M. Arif, T. Brouard and N. Vincent, *A fusion methodology for recognition of offline signatures*, in Proceedings of 4th International Workshop Pattern Recognition and Information System, pp. 35–44, 2004.

[37] L. Bovino, S. Impedovo, G. Pirlo and L. Sarcinella, *Multi-expert verification of handwritten signatures*, in Proceedings of 7th International Conference Document Analysis and Recognition, Edinburgh, U.K., pp. 932–936, 2003.

[38] M. Arif, T. Brouard, and N. Vincent, *A fusion methodology based on Dempster-Shafer evidence theory for two biometric applications*, in Proceedings of 18th International Conference on Pattern Recognition, Vol. 4, pp. 590-593, 2006.

[39] H. Cardot, M. Revenu, B. Victorri and M.J. Revillet, *A static signature verification system based on a cooperating neural networks architecture*, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 8(3), pp. 679–692, 1994.

[40] R. Bajaj and S. Chaudhury, Signature verification using multiple neural classifiers, Pattern Recognition, Vol. 30(1), pp. 1–7, 1997.

[41] Y. Hongo, D. Muramatsu and T. Matsumoto, *AdaBoost-based on-line signature verifier*, in Biometric Technology for Human Identification II, A.K. Jain and N.K. Ratha, Eds. Proc. SPIE, Vol. 5779, pp. 373–380, 2005.

[42] D. Muramatsu, K. Yasuda and T. Matsumoto, *Biometric Person Authentication Method Using Camera-Based Online Signature*, in Proceedings of 10th International Conference on Document Analysis and Recognition, Barcelona, Spain, pp. 46-50, July 2009.

[43] L. Wan, Z. Lin and R.C. Zhao, *Signature verification using integrated classifiers*, in the 4th Chinese Conference on Biometric Recognition, Beijing, China, pp. 7–8, 2003.

[44] R.L. Larkins, *Off-line Signature Verification*, Thesis of University of Waikato, 2009.

[45] K. Franke, L.R.B. Schomaker, C. Veenhuis, C. Taubenheim, I. Guyon, L.G. Vuurpijl, M. van Erp and G. Zwarts, *WANDA: A generic framework applied in forensic handwriting analysis and writer identification*, Design and Application of Hybrid Intelligent Systems, in Proceedings of 3rd International Conference on Hybrid Intelligent Systems, Abraham, A., Koeppen, M., & Franke, K., eds., IOS Press, Amsterdam, pp. 927-938, 2003.

[46] Ink Markup Language (InkML), W3C Working Draft 23 October 2006, http://www.w3.org/TR/InkML/#orientation.

[47] E.J. Candès, *Ridgelets: Theory and Applications*, Ph.D. thesis, Department of Statistics, Stanford University, 1998.

[48] S.G. Mallat, *A theory for multiresolution signal decomposition: The wavelet representation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11(7), pp. 674-693, 1989.

[49] N. Abbas and Y. Chibani, *Combination of Off-Line and On-Line Signature Verification Systems Based on SVM and DST*, in the 11th International Conference on Intelligent Systems Design and Applications, pp. 855–860, 2011.

[50] N. Abbas, and Y. Chibani, *SVM-DSmT combination for the simultaneous verification of off-line and on-line handwritten signatures*, International Journal of Computational Intelligence and Applications, Vol. 11(3), 2012.

[51] E.J.R. Justino, F. Bortolozzi and R. Sabourin, *A comparison of SVM and HMM classifiers in the off-line signature verification*, Pattern Recognition Letters, Vol. 26, pp. 1377-1385, 2005.

[52] V.N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 1995.

[53] H.P. Huang and Y.H. Liu, *Fuzzy support vector machines for pattern recognition and data mining*, International Journal of Fuzzy Systems, Vol. 4(3), pp. 826-835, 2002.

[54] N. Abbas and Y. Chibani, *SVM-DSmT Combination for Off-Line Signature Verification*, in the International Conference on Computer, Information and Telecommunication Systems, Amman, Jordan, 2012.

[55] A. Appriou, *Probabilités et incertitude en fusion de données multisenseurs*, Revue Scientifique et Technique de la Défense, Vol. 11, pp. 27-40, 1991.

[56] M. Kalera, B. Zhang and S. Srihari, *Offline Signature Verification and Identification Using Distance Statistics*, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 18(7), pp. 1339–1360, 2004.

[57] C.E. van den Heuvel, K. Franke, L. Vuurpijl (et al), *The ICDAR 2009 signature verification competition*, In ICDAR 2009 proceedings.

[58] http://www.sigcomp09.arsforensica.org, April 2009.

Automatic Aircraft Recognition using DSmT and HMM

Xin-de Li
Jin-dong Pan
Jean Dezert

Originally published as Li X.-D., Pan J.-D., Dezert J., A*utomatic Aircraft Recognition using DSmT and HMM*, in Proc. of Fusion 2014 Int Conf onInformation Fusion, Salamanca, Spain, July 7-10, 2014, and reprinted with permission.

***Abstract*—In this paper we propose a new method for solving the Automatic Aircraft Recognition (AAR) problem from a sequence of images of an unknown observed aircraft. Our method exploits the knowledge extracted from a training image data set (a set of binary images of different aircrafts observed under three different poses) with the fusion of information of multiple features drawn from the image sequence using Dezert-Smarandache Theory (DSmT) coupled with Hidden Markov Models (HMM). The first step of the method consists for each image of the observed aircraft to compute both Hu's moment invariants (the first features vector) and the partial singular values of the outline of the aircraft (the second features vector). In the second step, we use a probabilistic neural network (PNN) based on the training image dataset to construct the conditional basic belief assignments (BBA's) of the unknown aircraft type within the set of a predefined possible target types given the features vectors and pose condition. The BBA's are then combined altogether by the Proportional Conflict Redistribution rule #5 (PCR5) of DSmT to get a global BBA about the target type under a given pose hypothesis. These sequential BBA's give initial recognition results that feed a HMM-based classifier for automatically recognizing the aircraft in a multiple poses context. The last part of this paper shows the effectiveness of this new Sequential Multiple-Features Automatic Target Recognition (SMF-ATR) method with realistic simulation results. This method is compliant with real-time processing requirement for advanced AAR systems.**

Keywords: Information fusion; DSmT; ATR; HMM.

I. Introduction

ATR (Automatic Target Recognition) systems play a major role in modern battlefield for automatic monitoring and detection, identification and for precision guided weapon as well. The Automatic Aircraft Recognition (AAR) problem is a subclass of the ATR problem. Many scholars have made extensive explorations for solving ATR and AAR problems. The ATR method is usually based on target recognition using template matching [1], [2] and single feature (SF) extraction [3]–[7] algorithms. Unfortunately, erroneous recognition often occurs when utilizing target recognition algorithms based on single feature only, specially if there exist important changes in pose and appearance of aircrafts during flight path in the image sequence. In such condition, the informational content drawn from single feature measures cannot help enough to make a reliable classification. To overcome this serious drawback, new ATR algorithms based on multiple features (MF) and fusion techniques have been proposed [8]–[12]. An interesting MF-ATR algorithm based on Back-Propagation Neural Network (BP-NN), and Dempster-Shafer Theory (DST) of evidence [23] has been proposed by Yang et al. in [11] which has been partly the source of inspiration to develop our new improved sequential MF-ATR method presented here and introduced briefly in [12] (in chinese). In this paper we will explain in details how our new SMF-ATR method works and we evaluate its performances on a typical real image sequence.

Although MF-ATR approach reduces the deficiency of SF-ATR approach in general, the recognition results can sometimes still be indeterminate form a single image exploitation because the pose and appearance of different kinds of aircrafts can be very similar for some instantaneous poses and appearances. To eliminate (or reduce) uncertainty and improve the classification, it is necessary to exploit a sequence of images of the observed aircraft during its flight and develop efficient techniques of sequential information fusion for advanced (sequential) MF-ATR systems. Two pioneer works on sequential ATR algorithms using belief functions (BF) have been proposed in last years. In 2006, Huang et al. in [13] have developed a sequential ATR based on BF, Hu's moment invariants (for image features vector), a BP-NN for pattern classification, and a modified Dempster-Shafer (DS) fusion rule[1]. A SF-ATR approach using BF, Hu's moment invariants, BP-NN and DSmT rule has also been proposed in [14] the same year. In these papers, the authors did clearly show the benefit of the integration of temporal SF measures for the target recognition, but the performances obtained were still limited because of large possible changes in poses and appearances of observed aircrafts (specially in high maneuver modes as far as military aircrafts are under concern). The purpose of this paper is to develop a new (sequential) MF-ATR method able to provide a high recognition rate with a good robustness when face to large changes of poses and ppearances of observed aircraft during its flight.

The general principle of our SMF-ATR method is shown on Fig.1. The upper part of Fig. 1 consists in Steps 1 & 2, whereas the lower part of Fig. 1 consists in Steps 3 & 4 respectively described as follows:

- Step 1 (Features extraction) : We consider and extract only two features vectors in this work[2] (Hu's moment

[1] called the abortion method by the authors.
[2] The introduction of extra features is possible and under investigations.

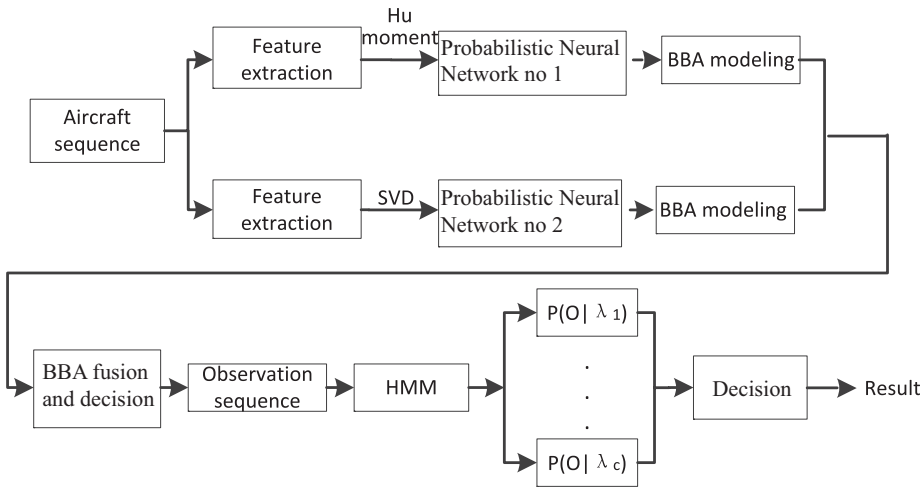


Fig. 1: General principle of our sequential MF-ATR approach.

invariants vector, and Singular Values Decomposition (SVD) features vector) from the binary images[3]

- Step 2 (BBA's construction[4]) : For every image in the sequence and from their two features vectors, two Bayesian BBA's on possible (target type,target pose) are computed from the results of two PNN's trained on the image dataset. The method of BBA construction is different from the one proposed in [12].
- Step 3 (BBA's combination) : For every image, say the k-th image, in the sequence, the two BBA's of Step 2 are combined with the PCR5 fusion rule, from which a decision O_k on the most likely target type and pose is drawn.
- Step 4 (HMM-based classifier) : From the sequence $O^K = \{O_1, \ldots, O_k \ldots, O_K\}$ of K local decisions computed at Step 3, we feed several HMM-based classifiers in parallel (each HMM characterizes each target type) and we find finally the most likely target observed in the image sequence which gives the output of our SMF-ATR approach.

The next section presents each step of this new SMF-ATR approach. Section 3 evaluates the performances of this new method on real image datasets. Conclusions and perspectives of this work are given in Section 4.

II. The sequential MF-ATR approach

In this section we present the aforementioned steps necessary for the implementation of our new SMF-ATR method.

[3]In this work, we use only with binary images because our image training dataset contains only binary images with clean backgrounds, and working with binary images is easier to do and requires less computational burden than working with grey-level or color images. Hence it helps to satisfy real-time processing. The binarization of the images of the sequence under analysis is done with the the Flood Fill Method explained in details in [22] using the point of the background as a seed for the method.

[4]The mathematical definition of a BBA is given in Section II-C.

A. Step 1: Features extraction from binary image

Because Aircraft poses in a flight can vary greatly, we need image features that are stable and remain unchanged under translation, rotation and scaling. In terms of aircraft features, two categories are widely used: 1) moment features and 2) contour features. Image moments have been widely used since a long time specially for pattern-recognition applications [16]. Moment features which are the descriptions of image regional characteristics are mainly obtained from the intensity of each pixel of target image. Contour features are extracted primarily by discretizing the outline contour and they describe the characteristic of the outline of the object in the image. In terms of moment features, Hu's moment invariants [6] are used here. As contour features, we use the SVD [15] of outlines extracted from the binary images.

• Hu's moments

Two-dimensional $(p+q)$-th order moments for $p, q = 0, 1, 2, ...$ of an image of size $M \times N$ are defined as follows:

$$m_{pq} \triangleq \sum_{m=1}^{M} \sum_{n=1}^{N} m^p n^q f(m,n) \tag{1}$$

where $f(m,n)$ is the value of the pixel (m,n) of the binary image. Note that m_{pq} may not be invariant when $f(m,n)$ by translation, rotating or scaling. The invariant features can be obtained using the $(p+q)$-th order central moments μ_{pq} for $p, q = 0, 1, 2, ...$ defined by

$$\mu_{pq} \triangleq \sum_{m=1}^{M} \sum_{n=1}^{N} (m-\bar{x})^p (n-\bar{y})^q f(m,n) \tag{2}$$

where $\bar{x}$, and $\bar{y}$ are the barycentric coordinates of image (i.e. the centroid of the image). These values are computed by $\bar{x} = \frac{m_{10}}{m_{00}} = \frac{1}{C}\sum_{m=1}^{M}\sum_{n=1}^{N} m \times f(m,n)$ and $\bar{y} = \frac{m_{01}}{m_{00}} = \frac{1}{C}\sum_{m=1}^{M}\sum_{n=1}^{N} n \times f(m,n)$, where C is a normalization constant given by $C = m_{00} = \sum_{m=1}^{M}\sum_{n=1}^{N} f(m,n)$. The centroid moments μ_{pq} is equivalent to the m_{pq} moment whose

center has been shifted to the centroid of the image. Therefore, μ_{pq} are invariant to image translations. Scale invariance is obtained by normalization [6]. The normalized central moments η_{pq} are defined for $p+q=2,3,\ldots$ by $\eta_{pq} \triangleq \mu_{pq}/\mu_{00}^{\gamma}$, with $\gamma=(p+q+2)/2$. Based on these normalized central moments Hu in [16] derived seven moment invariants that are unchanged under image scaling, translation and rotation as follows

$$\Phi_1 \triangleq \eta_{20}+\eta_{02}$$
$$\Phi_2 \triangleq (\eta_{20}-\eta_{02})^2+4\eta_{11}^2$$
$$\Phi_3 \triangleq (\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\eta_{03})^2$$
$$\Phi_4 \triangleq (\eta_{30}+\eta_{12})^2+(\eta_{21}+\eta_{03})^2$$
$$\Phi_5 \triangleq (\eta_{30}-3\eta_{12})(\eta_{30}+\eta_{12})[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2] + (3\eta_{21}-\eta_{03})(\eta_{21}+\eta_{03})[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2]$$
$$\Phi_6 \triangleq (\eta_{20}-\eta_{02})[(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2] + 4\eta_{11}(\eta_{30}+\eta_{12})(\eta_{21}+\eta_{03})$$
$$\Phi_7 \triangleq (3\eta_{21}-\eta_{03})(\eta_{30}+\eta_{12})[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2] - (\eta_{30}-3\eta_{12})(\eta_{21}+\eta_{03})[3(\eta_{30}+\eta_{12})^2-(\eta_{03}+\eta_{21})^2]$$

In this work, we use only the four simplest Hu's moments to compute, that is $\mathbf{\Phi}=[\Phi_1\ \Phi_2\ \Phi_3\ \Phi_4]$, to feed the first PNN of our sequential MF-ATR method[5].

• SVD features of the target outline

The SVD is widely applied signal and image processing because it is an efficient tool to solve problems with least squares method [21]. The SVD theorem states that if $\mathbf{A}_{m\times n}$ with $m>n$ (representing in our context the original binary data) is a real matrix[6], then it can be written using a so-called singular value decomposition of the form

$$\mathbf{A}_{m\times n}=\mathbf{U}_{m\times m}\mathbf{S}_{m\times n}\mathbf{V}_{n\times n}^T$$

where $\mathbf{U}_{m\times m}$ and $\mathbf{V}_{n\times n}$ are orthogonal[7] matrices. The columns of $\mathbf{U}$ are the left singular vectors. $\mathbf{V}^T$ has rows that are the right singular vectors. The real matrix $\mathbf{S}$ has the same dimensions as $\mathbf{A}$ and has the form[8]

$$\mathbf{S}_{m\times n}=\begin{bmatrix}\mathbf{S}_{r\times r} & \mathbf{0}_{r\times(n-r)}\\ \mathbf{0}_{r\times(m-r)} & \mathbf{0}_{(m-r)\times(n-r)}\end{bmatrix}$$

where $\mathbf{S}_{r\times r}=Diag\{\sigma_1,\sigma_2,\ldots,\sigma_r\}$ with $\sigma_1\geq\sigma_2,\geq\ldots\geq\sigma_r>0$ and $1\leq r\leq\min(m,n)$.

Calculating the SVD consists of finding the eigenvalues and eigenvectors of $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$. The eigenvectors of $\mathbf{A}^T\mathbf{A}$ make up the columns of $\mathbf{V}$, the eigenvectors of $\mathbf{A}\mathbf{A}^T$ make up the columns of $\mathbf{U}$. The singular values σ_1,..., σ_r are the diagonal entries of $\mathbf{S}_{r\times r}$ arranged in descending order, and they are square roots of eigenvalues from $\mathbf{A}\mathbf{A}^T$ or $\mathbf{A}^T\mathbf{A}$.

A method to calculate the set of discrete points $\{a_1,a_2,\ldots,a_n\}$ of a target outline from a binary image is proposed in [17]. The SVD features are then computed from the eigenvalues of the circulant matrix built from the discretized shape of the outline characterized by the vector $\mathbf{d}=[d_1,d_2,\ldots,d_n]$ where d_i is the distance of the centroid of the outline to the discrete points a_i, $i=1,2,\ldots,n$ of the outline.

In our analysis, it has been verified from our image dataset that only the first components of SVD features vector $\boldsymbol{\sigma}=[\sigma_1,\sigma_2,\ldots,\sigma_r]$ take important values with respect to the other ones. The other components of $\boldsymbol{\sigma}$ tend quickly towards zero. Therefore only few first components of $\boldsymbol{\sigma}$ play an important role to characterize the main features of target outline. However, if one considers only these few main first components of $\boldsymbol{\sigma}$, one fails to characterize efficiently some specific features (details) of the target profile. By doing so, one would limit the performances of ATR. That is why we propose to use the partial SVDs of outline as explained in the next paragraph.

To capture more details of aircraft outline with SVD, one has to taken into account also additional small singular values of SVD. This is done with the following procedure issued from the face recognition research community [24]. The normalized distance vector $\tilde{\mathbf{d}}=[\tilde{d}_1,\tilde{d}_2,\ldots,\tilde{d}_n]$ is built from $\mathbf{d}$ by taking $\tilde{\mathbf{d}}=[1,d_2/d_1,\ldots,d_n/d_1]$, where d_1 is the distance between the centroid of outline and the first chosen points of the contour of the outline obtained by a classical[9] edge detector algorithm. To capture the details of target outline and to reduce the computational burden, one works with partial SVDs of the original outline by considering only l sliding sub-vectors $\tilde{\mathbf{d}}_w$ of $\tilde{\mathbf{d}}$, where w is the number of components of $\tilde{\mathbf{d}}_w$. For example if one takes $w=3$ points only in the sub-vectors and if $\tilde{\mathbf{d}}=[\tilde{d}_1,\tilde{d}_2,\ldots,\tilde{d}_9]$, then one will take the sub-vectors $\tilde{\mathbf{d}}_w^1=[\tilde{d}_1,\tilde{d}_2,\tilde{d}_3]$, $\tilde{\mathbf{d}}_w^2=[\tilde{d}_4,\tilde{d}_5,\tilde{d}_6]$ and $\tilde{\mathbf{d}}_w^3=[\tilde{d}_7,\tilde{d}_8,\tilde{d}_9]$ if we don't use overlapping components between sub-vectors. From the sub-vectors, one constructs their corresponding circulant matrix and apply their SVD to get partial SVD features vectors $\boldsymbol{\sigma}_w^{l=1}$, $\boldsymbol{\sigma}_w^{l=2}$, etc. The number l of partial SVD of the original outline of the target is given by $l=(n-w)/(w-m)+1$, where m is the number of components overlapped by each two adjacent sub-vectors, and n is the total number of discrete contour points of the outline given by the edge detector.

B. Step 2: BBA's construction with PNN's

In order to exploit efficiently fusion rules dealing with conflicting information modeled by belief mass assignments (BBA's) [18], [23], we need to build BBA's from all features computed from images of the sequence under analysis. The construction of the BBA's needs expert knowledge or knowledge drawn from training using image dataset. In this paper, we propose to utilize probabilistic neural networks (PNN) initially developed in nineties by Specht [19] to construct the BBA's because it is a common technique used in the target recognition and pattern classification community that is able to

[5]It is theoretically possible to work with all seven Hu's moments in our MF-ATR method, but we did not test this yet in our simulations.

[6]For a complex matrix $\mathbf{A}$, the singular value decomposition is $\mathbf{A}=\mathbf{U}\mathbf{S}\mathbf{V}^H$, where $\mathbf{V}^H$ is the conjugate transpose of $\mathbf{V}$.

[7]They verify $\mathbf{U}_{m\times m}^T\mathbf{U}_{m\times m}=\mathbf{I}_{m\times}$ and $\mathbf{V}_{n\times n}^T\mathbf{V}_{n\times n}=\mathbf{I}_{n\times n}$, where $\mathbf{I}_{m\times m}$ and $\mathbf{I}_{n\times n}$ are respectively the identity matrices of dimensions $m\times m$ and $n\times n$.

[8]$\mathbf{0}_{p\times q}$ is a $p\times q$ matrix whose all its elements are zero.

[9]In this work, we use the *cvcontour* function of opencv software [22] to extract the target outline from a binary image.

achieve with large training dataset performances close to those obtained by a human expert in the field. The details of PNN's settings for BBA's construction are given in [12]. However, because the neural network after training to some extent has a good discriminant ability (close to an expert in the field), the BBA is constructed by the neural network directly based on the PNN's output, which is different from the construction of the BBA based on the confusion matrix described in [12].

Here we present how the two PNN's (shown in Figure 1) work. In our application, we have $N_c = 7$ types of aircrafts in our training image dataset. For each type, the aircraft is observed with $N_p = 3$ poses. Therefore we have $N_{cp} = N_c \times N_p = 21$ types of distinct cases in our dataset. For each case, one has $N_i = 30$ images available for the training. Therefore the whole training dataset contains $N_{cpi} = N_c N_p N_i = 7 \times 3 \times 30 = 630$ binary images. For the first PNN (fed by Hu's features vector), the number of input layer neurons is 4 because we use only $\mathbf{\Phi} = [\Phi_1, \Phi_2, \Phi_3, \Phi_4]$ Hu's moment invariants in this work. For the second PNN (fed by partial SVD features vector), the number of input layer neurons is constant and equal to $l \times w$ because we take l windows with the width w (so one has w singular values of partial SVD for every window). The number of hidden layer neurons of each PNN is the number of the training samples, $N_{cpi} = 630$. The number of output layer neurons is equal to $N_{cp} = 21$ (the number of different possible cases).

Our PNN's fed by features input vectors (Hu's moments and SVD outline) do not provide a hard decision on the type and pose of the observed target under analysis because in our belief-based approach we need to build BBA's. Therefore the competition function of the output layer for decision-making implemented classically in the PNN scheme is not used in the exploitation[10] phase of our approach. Instead, the PNN computes the $N_{cp} \times N_i$ (Euclidean) distances between the features vectors of the image under test and the $N_{cpi} = 630$ features vectors of the training dataset. A Gaussian radial basis function (G-RBF) is used in the hidden layer of the PNN's [19] to transform its input (Euclidean) distance vector of size $1 \times N_{cpi}$ into another $1 \times N_{cpi}$ distance (similarity) that feeds the output layer through a weighting matrix of size $N_{cpi} \times N_{cp} = 630 \times 21$ estimated from the training samples. As a final output of each PNN, we get an unnormalized similarity vector $\mathbf{m}$ of size $(1 \times N_{cpi}) \times (N_{cpi} \times N_{cp}) = 1 \times N_{cp} = 1 \times 21$ which is then normalized to get a Bayesian BBA on the frame of discernment $\Theta = \{(target_i, pose_j), i = 1, \ldots, c, j = 1, \ldots, p\}$. Because we use only two[11] PNN's in this approach, we are able to build two Bayesian BBA's $m_1(.)$ and $m_2(.)$ defined on the same frame Θ for every image of the sequence to analyze.

C. Step 3: Fusion of BBA's and local decision

A basic belief assignment (BBA), also called a (belief) mass function, is a mapping $m(.) : 2^\Theta \mapsto [0; 1]$ such that $m(\emptyset) = 0$ and $\sum_{X \in 2^\Theta} m(X) = 1$, where Θ is the so-called frame of discernment of the problem under concern which consists of a finite discrete set of exhaustive and exclusive hypotheses[12] θ_i, $i = 1, \ldots, n$, and where 2^Θ is the power-set of Θ (the set of all subsets of Θ). This definition of BBA has been introduced in Dempster-Shafer Theory (DST) [23]. The focal elements of a BBA are all elements X of 2^Θ such that $m(X) > 0$. Bayesian BBA's are special BBA's having only singletons (i.e. the elements of Θ) as focal elements.

In DST, the combination of BBA's is done by Dempster's rule of combination [23] which corresponds to the normalized conjunctive consensus operator. Because this fusion rule is known to be not so efficient (both in highly and also in low conflicting) in some practical situations [25], many alternative rules have been proposed during last decades [18], Vol. 2.

To overcome the practical limitations of Shafers' model and in order to deal with fuzzy hypotheses of the frame, Dezert and Smarandache have proposed the possibility to work with BBA's defined on Dedekind's lattice[13] D^Θ [18] (Vol.1) so that intersections (conjunctions) of elements of the frame can be allowed in the fusion process, with eventually some given restrictions (integrity constraints). Dezert and Smarandache have also proposed several rules of combination based on different Proportional Conflict Redistribution (PCR) principles. Among these new rules, the PCR5 and PCR6 rules play a major role because they do not degrade the specificity of the fusion result (contrariwise to most other alternative rule), and they preserve the neutrality of the vacuous BBA[14]. PCR5 and PCR6 provide same combined BBA when combining only two BBA's $m_1(.)$ and $m_2(.)$, but they differ when combining three (or more) BBA's altogether. It has been recently proved in [26] that PCR6 is consistent with empirical (frequentist) estimation of probability measure, unlike other fusion rules[15].These two major differences with DST, make the basis of Dezert-Smarandache Theory (DSmT) [18].

In the context of this work, we propose to use PCR5 to combine the two (Bayesian) BBA's $m_1(.)$ and $m_2(.)$ built from the two PNN's fed by Hu's features vector and SVD outline features vector. Because for each image of the observed target in the sequence, one has only two BBA's to combine, the PCR5 fusion result is same as the PCR6 fusion result. Of course, if one wants to include other kinds of features vectors with additional PNN's, the PCR6 fusion rule is recommended. The PCR principle consists in redistributing the partial conflicting masses[16]only to the sets involved in the conflict and proportionally to their mass. The PCR5 (or PCR6) combination of

[10] when analyzing a new sequence of an unknown observed aircraft.

[11] A first PPN fed by Hu's features, and a second PNN fed by SVD outline features – see Fig. 1.

[12] This is what is called Shafer's model of the frame in the literature.

[13] Dedekind's lattice is the set of all composite subsets built from elements of Θ with $\cup$ and $\cap$ operators.

[14] A vacuous BBA is the BBA such that $m(\Theta) = 1$.

[15] except the averaging rule.

[16] For two BBA's, a partial conflicting mass is a product $m_1(X)m_2(Y) > 0$ of the element $X \cap Y$ which is conflicting, that is such that $X \cap Y = \emptyset$.

two BBA's is done according to the following formula[17] [18]

$$m_{PCR5/6}(X) = \sum_{\substack{X_1,X_2\in 2^\Theta \\ X_1\cap X_2 = X}} m_1(X_1)m_2(X_2) + \sum_{\substack{Y\in 2^\Theta\setminus\{X\} \\ X\cap Y=\emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X)+m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X)+m_1(Y)}\right] \quad (3)$$

where all denominators in (3) are different from zero, and $m_{PCR5/6}(\emptyset) = 0$. If a denominator is zero, that fraction is discarded. All propositions/sets are in a canonical form. Because we work here only with Bayesian BBA's, the previous fusion formula is in fact rather easy to implement, see [18] (Vol. 2, Chap. 4).

In summary, the target features extraction in a sequence of K images allows us to generate, after Step 3, a set of BBA's $\{m^{Image_k}(.), k = 1, 2, \ldots, K\}$. Every BBA $m^{Image_k}(.)$ is obtained by the PCR5/6 fusion of BBA's $m_1^{Image_k}(.)$ and $m_2^{Image_k}(.)$ built from the outputs of two PNN's. From this combined BBA, a local[18] decision O_k can be drawn about the target type and target pose in $Image_k$ by taking the focal element of $m^{Image_k}(.)$ having the maximum mass of belief.

D. Step 4: Hidden Markov Model (HMM) for recognition

Usually (and specially in military context), the posture of an aircraft can continuously change a lot during its flightpath making target recognition based only on single observation (image) very difficult, because some ambiguities can occur between extracted features with those stored in the training image data set. To improve the target recognition performance and robustness, one proposes to use the sequence of target recognition decision O_k drawn from BBA's $\{m^{Image_k}(.), k = 1, 2, \ldots, K\}$ to feed HMM classifiers in parallel. We suggest this approach because the use of HMM has already been proved to be very efficient in speech recognition, natural language and face recognition. We briefly present HMM, and then we will explain how HMMs are used for automatic aircraft recognition.

Let us consider a dynamical system with a finite set of possible states $S = \{s_1, s_2, \ldots, s_N\}$. The state transitions of the system is modeled by a first order Markov chain governed by the transition probabilities given by $P(s(t_k) = s_j | s(t_{k-1}) = s_i, s(t_{k-2}) = s_k, \ldots) = P(s(t_k) = s_j | s(t_{k-1}) = s_i) = a_{ij}$, where $s(t_k)$ is the random state of the system at time t_k. A HMM is a doubly stochastic processes including an underlying stochastic process (i.e. a Markov chain for modeling the state transitions of the system), and a second stochastic process for modeling the observation of the system (which is a function of the random states of the system). A HMM, denoted $\lambda = (\mathbf{A}, \mathbf{B}, \Pi)$, is fully characterized by the knowledge of the following parameters

1) The number N of possible states $S = \{s_1, s_2, \ldots, s_N\}$ of the Markov chain.
2) The state transition probability matrix[19] $\mathbf{A} = [a_{ij}]$ of size $N \times N$, where $a_{ij} \triangleq P(s(t_k) = s_i | s(t_{k-1}) = s_j)$.
3) The prior mass function (pmf) Π of the initial state of the chain, that is $\Pi = \{\pi_1, \ldots, \pi_N\}$ with $\sum_{i=1}^N \pi_i = 1$, where $\pi_i = P(s(t_1) = s_i)$.
4) The number M of possible values $V = \{v_1, \ldots, v_M\}$ taken by the observation of the system.
5) The conditional pmfs of observed values given the states of the system characterized by the matrix $\mathbf{B} = [b_{mi}]$ of size $M \times N$, with $b_{mi} \triangleq P(O_k = v_m | s(t_k) = s_i)$, where O_k is the observation of the system (i.e. the local decision on target type with its pose) at time t_k.

In this work we consider a set of N_c HMMs in parallel, where each HMM is associated with a given type of target to recognize. We consider the following state and observation models in our HMMs:

- **State model:** For a given type of aircraft, we consider a finite set of distinct aircraft postures available in our training image dataset. In our application, we consider only three states corresponding to s_1 = *top view*, s_2 = *side view* and s_3 = *front view* as shown (for a particular aircraft) in Figure 2.

Fig. 2: Example of HMM states.

- **Observation model:** In our HMMs, we assume that each state (posture) of aircraft is observable. Since we have only $N_p = 3$ states $S = \{s_1, s_2, s_3\}$ for each aircraft, and we have $N_c = 7$ types of aircrafts in the training dataset, we have to deal with $N_{cp} = 3 \times 7 = 21$ possible[20] observations (local decisions) at each time t_k. As explained previously, at the end of Step 3 we have a set of BBA's $\{m^{Image_k}(.), k = 1, 2, ..., K\}$ that helps to draw the sequence of local decisions $O^K \triangleq \{O_1, \ldots, O_k, \ldots, O_K\}$. This sequence of decisions (called also *recognition observations*) is used to evaluate the likelihood $P(O^K|\lambda_i)$ of the different HMMs described by the parameter $\lambda_i = (\mathbf{A}_i, \mathbf{B}_i, \Pi_i)$, $i = 1, 2, \ldots, N_c$. The computation of these likelihoods will be detailed at the end of this section. The final decision for ATR consists to infer the true target type based on the maximum likelihood criterion. More precisely, one will decide that the target type is $i^\star$ if $i^\star = \arg\max_i P(O^K|\lambda_i)$.

• Estimation of HMM parameters

To make recognition with HMMs, we need at first to define a HMM for each type of target one wants to recognize. More precisely, we need to estimate the parameters $\lambda_i =$

[17]Here we assume that Shafers' model holds. The notation $m_{PCR5/6}$ means PCR5 and PCR6 are equivalent when combining two BBA's.

[18]because it is based only on a single image of the unknown observed target in the sequence under analysis.

[19]We assume that the transition matrix is known and time-invariant, i.e. all elements a_{ij} do not depend on t_{k-1} and t_k.

[20]We assume that the unknown observed target type belongs to the set of types of the dataset, as well as its pose.

$(\mathbf{A}_i, \mathbf{B}_i, \Pi_i)$, where $i = 1, \ldots, N_c$ is the target type in the training dataset. The estimation of HMM parameters is done from observation sequences drawn from the training dataset with Baum-Welch algorithm [20] that must be initialized with a chosen value $\lambda_i^0 = (\mathbf{A}_i^0, \mathbf{B}_i^0, \Pi_i^0)$. This initial value is chosen as follows:

1) – State prior probabilities Π_i^0 for a target of type i: For each HMM, we consider only three distinct postures (states) s_1, s_2 and s_3 for the aircraft. We use a uniform prior probability mass distribution for all types of targets. Therefore, we take $\Pi_i^0 = [1/3, 1/3, 1/3]$ for any target type $i = 1, \ldots, N_c$ to recognize.

2) – State transition matrix $\mathbf{A}_i^0$ of a target of type i: The components a_{pq} of the state transition matrix $\mathbf{A}_i^0$ are estimated from the analysis of many sequences[21] of target i as follows

$$a_{pq} = \frac{\sum_{k=1}^{K-1} \delta(s(t_k), s_p) \times \delta(s(t_{k+1}), s_q)}{\sum_{k=1}^{K-1} \delta(s(t_k), s_p)} \quad (4)$$

where N_p is the number of states of the Markov chain, $\delta(x, y)$ is the Kronecker delta function defined by $\delta(x, y) = 1$ if $y = x$, and $\delta(x, y) = 0$ otherwise, and where K is the number of images in the sequence of target i available in the training phase. For example, if in the training phase and for a target of type $i = 1$, we have the following sequence of (target type, pose) cases given by $[(1,1),(1,1),(1,2),(1,1),(1,3),(1,1),(1,1)]$, then from Eq. (4) with $K = 7$, we get[22]

$$\mathbf{A}_{i=1}^0 = \begin{bmatrix} 2/4 & 1/4 & 1/4 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

3) – Observation matrix $\mathbf{B}_i^0$ for a target of type i: The initial observation matrix $\mathbf{B}_i^0$ is given by the confusion matrix learnt from all images of the training dataset. More precisely, from every image of the training dataset, we extract Hu's features and partial SVD outline features and we feed each PNN to get two BBA's according to Steps 1-3. From the combined BBA, we make the local decision $(target_i, pose_j)$ if $m((target_i, pose_j))$ is bigger than all other masses of belief of the BBA. This procedure is applied to all images in the training dataset. By doing so, we can estimate empirically the probabilities to decide $(target_i, pose_j)$ when real case $(target_{i'}, pose_{j'})$ occurs. So we have an estimation of all components of the global confusion matrix $\mathbf{B}^0 = [P(decision = (target_i, pose_j) \mid reality = (target_{i'}, pose_{j'}))]$. From $\mathbf{B}^0$ we extract the c sub-matrices (conditional confusion matrices) $\mathbf{B}_i^0$, $i = 1, \ldots, N_c$ by taking all the rows of $\mathbf{B}^0$ corresponding to the target of type i. In our application, one has $N_c = 7$ types and $N_p = 3$ postures (states) for each target type, hence one has $N_{cp} = 7 \times 3 = 21$ possibles observations. Therefore the global confusion matrix $\mathbf{B}^0$ has size 21×21 is the stack of $N_c = 7$ sub-matrices $\mathbf{B}_i^0$, $i = 1, ..., N_c$, each of size $N_p \times N_{cp} = 3 \times 21$.

[21]The video stream of different (known) aircraft flights generate the sequences of images to estimate approximately a_{pq}

[22]One verifies that the probabilities of each raw of this matrix sum to 1.

• Exploitation of HMM for ATR

Given a sequence O^K of K local decisions drawn from the sequence of K images, and given N_c HMMs characterized by their parameter λ_i $(i = 1, \ldots, N_c)$, one has to compute all the likelihoods $P(O^K|\lambda_i)$, and then infer from them the true target type based on the maximum likelihood criterion which is done by deciding the target type $i^\star$ if $i^\star = \arg\max_i P(O^K|\lambda_i)$. The computation of $P(O^K|\lambda_i)$ is done as follows [20]:

- generation of all possible state sequences of length K, $S_l^K = [s_l(t_1)s_l(t_2)\ldots s_l(t_K)]$, where $s_l(t_k) \in S$ (k=1,..., K) and $l = 1, 2, \ldots, |S|^K$
- computation of $P(O^K|\lambda_i)$ by applying the total probability theorem as follows [23]

$$P(S_l^K|\lambda_i) = \pi_{s_l(t_1)} \cdot a_{s_l(t_1)s_l(t_2)} \cdot \ldots \cdot a_{s_l(t_{K-1})s_l(t_K)} \quad (5)$$

$$P(O^K|\lambda_i, S_l^K) = b_{s_l(t_1)O_1} \cdot b_{s_l(t_2)O_2} \cdot \ldots \cdot b_{s_l(t_K)O_K} \quad (6)$$

$$P(O^K|\lambda_i) = \sum_{l=1}^{|S|^K} P(O^K|\lambda_i, S_l^K) P(S_l^K|\lambda_i) \quad (7)$$

III. Simulations results

For the simulations of SMF-ATR method, we have used $N_c = 7$ types of aircrafts in the training image dataset. Each image of the sequence has 1200×702 pixels. The sequences of aircraft observations in the training dataset take 150 frames. The $N_p = 3$ poses of every aircraft is shown in Fig. 3. For evaluating our approach, we have used sequences (test samples) of images of 7 different aircraft, more precisely the Lockheed-F22, Junkers-G.38ce, Tupolev ANT 20 Maxime Gorky, Caspian Sea Monster (Kaspian Monster), Mirage-F1, Piaggio P180, and Lockheed-Vega, flying under conditions that generate a lot of state (posture) changes in the images. The number of the images in each sequence to test varies from 400 to 500. The shaping parameter of the G-RBF of PNN's has been set to 0.1. The simulation is done in two phases: 1) the training phase (for training PNN's and estimating HMM's parameters), and 2) the exploitation phase for testing the real performances of the SMF-ATR with test sequences.

A - Performances evaluation

In our simulations, we have tested SMF-ATR with two different fusion rules: 1) the PCR5 rule (see Section II-C), and 2) Dempster-Shafer (DS) rule[24] [23]. The percentages of successful recognition (i.e. the recognition rate R_i) obtained with these two SMF-ATR methods are shown in Table I for each type $i = 1, 2, \ldots, N_c$ of aircraft. The performances of these SMF-ATR versions are globally very good since one is able to recognize with a minimum of 85.2% of success the types of aircraft included in the image sequences under test when using DS-based SMF-ATR, and with a minimum of

[23]The index i of components of $\mathbf{A}_i$ and $\mathbf{B}_i$ matrices has been omitted for notation convenience in the last two formulas.

[24]Because Dempster's rule is one of the basis of Dempster-Shafer Theory, we call prefer to call it Dempster-Shafer rule, or just DS rule. This rule coincides here with Bayesian fusion rule because we combine two Bayesian BBA's and we don't use informative priors.

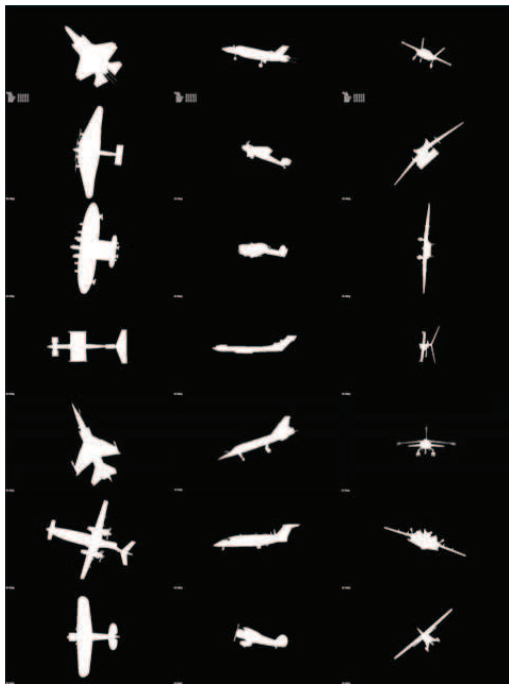


Fig. 3: Poses of different types of aircrafts.

93.5% of success with the PCR5-based SMF-ATR. In term of computational time, it takes between 5ms and 6ms to process each image in the sequence with no particular optimization of our simulation code, which indicates that this SMF-ATR approach is close to meet the requirement for real-time aircraft recognition. It can be observed that PCR5-based SMF-ATR outperforms DS-based SMF-ATR for 3 types of aircraft and gives similar recognition rate as with DS-based SMF-ATR for other types. So PCR5-based SMF-ATR is globally better than DS-based SMF-ATR for our application.

Target type	1	2	3	4	5	6	7
R_i (PCR5 rule)	95.7	93.5	96.3	98.2	96.3	98.5	97.3
R_i (DS rule)	95.7	93.5	85.2	97.8	96.3	98.5	97.2

TABLE I: Aircraft recognition rates R_i (in %).

B - Robustness of SMF-ATR to image scaling

To evaluate the robustness of (PCR5-based) SMF-ATR approach to image scaling effects, we did apply scaling changes (zoom out) of $ZO = 1/2$, $ZO = 1/4$ and $ZO = 1/8$ in the images of the sequences under test. The performances of the SMF-ATR are shown in Table II. One sees that the degradation of recognition performance of SMF-ATR due to scaling effects is very limited since even with a 1/8 zoom out one gets 90% of successful target recognition. The performance will decline sharply if the targets zoom out goes beyond 1/16.

C - Robustness to compound type

Table III gives the performances of SMF-ATR on sequences with two types of targets (475 images with type 1, and 382 images with type 2).

The two left columns of Table III show the performances

Target type	1	2	3	4	5	6	7
R_i (no ZO)	95.7	93.5	96.3	98.2	96.3	98.5	97.3
R_i (ZO=1/2)	95.0	92.0	95.2	94.7	96.1	96.6	95.4
R_i (ZO=1/4)	95.0	92.0	94.7	91.7	93.6	91.6	95.7
R_i (ZO=1/8)	95.0	92.2	93.1	89.3	93.6	94.5	90.7

TABLE II: Aircraft recognition rates R_i (in %) of (PCR5/6-based) SMF-ATR with different zoom out values.

Aircraft	Single Type 1	Single Type 2	Compound Type
R_i (SMF-ATR)	96.3 %	98.5%	97.3%

TABLE III: Robustness to target compound.

obtained when recognizing each type separately in each subsequence. The last column shows the performance when recognizing the compound type Type 1 $\cup$ Type 2. One sees that the performance obtained with compound type (97.3%) is close to the weighted average[25] 97.5% recognition rate. This indicates that no wide range of recognition errors occurs when the targets type change during the recognition process, making SMF-ATR robust to target type switch.

D - Performances with and without HMMs

We have also compared the performances of SMF-ATR, with two methods using more features but which do not exploit sequences of images with HMM. More precisely, the recognition is done locally from the combined BBA for every image without temporal integration processing based on HMM. We call these two Multiple Features Fusion methods MFF1 and MFF2 respectively. In MMF1, one uses Hu's moments, NMI (Normalized Moment of Inertia), affine invariant moments, and SVD of outline, PNN and PCR5 fusion, whereas MMF2 uses same features as MMF1 but with BP network as classifier and DS rule of combination. The recognition performances are shown in Table IV. One sees clearly the advantage to use the image sequence processing with HMMs because of significant improvement of ATR performances. The recognition rate of MFF2 declines seriously because the convergence of the BP network is not good enough.

Target type	1	2	3	4	5	6	7
R_i (SMF-ATR)	95.7	93.5	96.3	98.2	96.3	98.5	97.3
R_i (MFF1)	89.2	92.0	91.2	86.9	92.2	93.5	95.0
R_i (MFF2)	64.9	51.6	82.8	82.2	70.8	48.3	58.9

TABLE IV: Performances (in %) with and without HMMs.

E - SMF-ATR versus SSF-ATR

We have also compared in Table V the performances SMF-ATR with those of two simple SSF-ATR[26] methods, called SSF1-ATR and SSF2-ATR. The SSF1-ATR uses only Hu's moments features whereas SSF2-ATR uses only SVD of outline as features. SSF1-ATR exploits image sequence information using BP networks as classifier and DS rule for combination, while SSF2-ATR uses PNN and PCR5/6 rule.

[25] According to the proportion of the two types in the whole sequence.

[26] SSF-ATR stands for Single-feature Sequence Automatic Target Recognition.

Target type	1	2	3	4	5	6	7
R_i (SMF-ATR)	95.7	93.5	96.3	98.2	96.3	98.5	97.3
R_i (SFF1-ATR)	39.3	42.3	74.3	56.7	60.1	33.9	44.3
R_i (SFF2-ATR)	88.8	66.4	86.7	66.9	73.6	52.9	63.8

TABLE V: Performances (in %) of SMF-ATR and SFF-ATR.

One clearly sees the serious advantage of SMF-ATR with respect to SFF-ATR due to the combination of information drawn from both kinds of features (Hu's and SVD of outline) extracted from the images.

IV. Conclusions and perspectives

A new SMF-ATR approach based on features extraction has been proposed. The extracted features from binary images feed PNNs for building basic belief assignments that are combined with DSmT PCR rule to make a local (based on one image only) decision on target type. The set of local decisions acquired over time for the image sequence feeds HMMs to make the final recognition of the target. The evaluation of this new SMF-ATR approach has been done with realistic sequences of aircraft observations. SMF-ATR is able to achieve higher recognition rates than classical approaches that do not exploit HMMs, or SSF-ATR. Another complementary analysis of the robustness of SMF-ATR to target occultation is currently under progress and will be published in a forthcoming paper. Our very preliminary results based only on few sequences indicate that SMF-ATR seems very robust to target occultations occurring randomly in single (non consecutive) images, but a finer analysis based on Monte-Carlo simulation will be done to evaluate quantitatively its robustness in different conditions (number of consecutive occultations in the sequences, the level of occultation, etc). As interesting perspectives, we want to extend SMF-ATR approach for detecting new target types that are not included in image data set. Also, we would want to deal with the recognition of multiple crossing targets observed in a same image sequence.

Acknowledgment

This work was supported by NNSF of China (No. 60804063, 61175091), and Excellent Young Teacher Foundation sponsored by Qing Lan Project.

References

[1] G. Marsiglia, L. Fortunato, A. Ondini, G. Balzarotti, Template matching techniques for automatic IR target recognition in real and simulated scenarios: tests and evaluations, In Proc. of SPIE: Target Recognition XIII, Vol. 5094, pp. 159–169, 2003.

[2] O.K. Kwon, D.G. Sim, R.H. Park, Robust Hausdorff distance matching algorithms using pyramidal structures, Pattern Reco., Vol. 34 (10), pp. 2005–2013, 2001.

[3] S. Marouani, A. Huertas, G. Medioni, Model-based aircraft recognition in perspective aerial imagery, In Proc. of Int. Symp. on Computer Vision, pp. 371–376, Coral Gables, FL, USA, 1995.

[4] S. Das, B. Bhanu, A system for model-based object recognition in perspective aerial images, Pattern Reco., Vol. 31 (4), pp. 465–491, 1998.

[5] J.F. Gilmore, Knowledge-based target recognition system evolution, Optical Eng., Vol. 30 (5), pp. 557–570, 1991.

[6] S.A. Dudani, K.J. Breeding, R.B. McGhee, Aircraft identification by moment invariants, IEEE Trans. on Comp., Vol. 100, no. 1, 1977.

[7] Y.-J. Du, H.-C. Zhang, Q. Pan, Three-dimensional aircraft recognition using moments, Journal of Data Acquisition and Processing, Vol. 15 (3), pp. 390–394, 2000.

[8] L. Gu, Z.-Q. Zhuang, G.-Y. Zheng, Z.-J. Wang, Algorithm for hand shape matching based on feature fusion, Computer Applications, Vol. 25 (10), pp. 2286–2288, 2005.

[9] C.-Q. Deng, G. Feng, Content-based image re-trieval using combination features, Computer Applications, Vol. 23 (7), pp. 100–102, 2003.

[10] L. Chen, J. Chen, Multi-feature fusion method based on support vector machine and K-nearest neighbor classifier, Computer Applications, Vol. 29 (3), pp. 833–835, 2009.

[11] F.-P. Yang, Z.-X. Bai, Target Recognition Method based on Combination Of BP Neural Network with-Evidence Theory, Fire Control and Command Control, Vol. 31 (10), pp. 88–90, 2006.

[12] X.-D. Li, W.-D.Yang, J. Dezert, An Airplane Image Target's Multi-feature Fusion Recognition Method, Acta Automatica Sinica, Vol. 38 (8), pp. 1298–1307, 2012.

[13] J. Huang, Y. Liang, Y.-M. Cheng, Q. Pan, J.-W.Hu, Automatic target recognition method based on sequential images, J. of Computer Applications, Vol. 27 (1), pp. 87–93, 2006.

[14] J. Hou, Z. Miao, Q. Pan, Intelligent target recognition method of sequential images based on DSmT, J. of Computer Applications, Vol. 26 (1), pp. 120–122, 2006.

[15] S. Chen, J. Feng, Visual pattern recognition by moment invariants, in Proc. of 2010 IEEE Int. Conf. on Inform. and Autom., Harbin, China, pp. 857–860, 2010.

[16] M.K. Hu, Visual pattern recognition by moment invariants, IRE Trans. on Information Theory, Vol. IT-8 (2), pp. 179–187, 1962.

[17] J.-D. Pan, X.-D. Li, An airplane target recognition method based on singular value and PNN, Aero Weaponry, Vol. 1, pp. 45–50, 2013.

[18] F. Smarandache, J. Dezert, Advances and applications of DSmT for information fusion (Collected works), American Research Press, USA Vol.1, 2 & 3, 2004–2009. http://www.onera.fr/staff/jean-dezert?page=2

[19] D.F. Specht, Probabilistic neural networks, Neural networks, Vol. 3, No. 1, pp. 109–118, 1990.

[20] L.R. Rabiner, A tutorial on Hidden Markov Models and selected applications in speech recognition, Proc. of the IEEE, Vol. 77, No. 2, pp. 257–286, 1989. http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/

[21] J.C. Nash, The Singular-Value Decomposition and Its Use to Solve Least-Squares Problems, Ch. 3 in Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation, 2nd ed. Bristol, England: Adam Hilger, pp. 30–48, 1990.

[22] G. Bradski, A. Kaehler, Learning OpenCV: Computer vision with the OpenCV library, O'Reilly, pp. 124–129, 2008.

[23] G. Shafer, A Mathematical theory of evidence, Princeton University Press, Princeton, NJ, U.S.A., 1976.

[24] L. Cao Lin, D.-F. Wang, X.-J. Liu, M.-Y. Zou, Face recognition based on two-dimensional Gabor wavelets, Journal of Electronics & Information Technology, Vol. 28, No. 3, pp. 490–494, 2006.

[25] J. Dezert, A. Tchamova, On the validity of Dempster's fusion rule and its interpretation as a generalization of Bayesian fusion rule, Int. J. of Intel. Syst., Vol. 29, No. 3, Dec. , 2013.

[26] F. Smarandache, J. Dezert, On the consistency of PCR6 with the averaging rule and its application to probability estimation, in Proc. of Fusion 2013, Istanbul, Turkey, July 9-12, 2013.

Static versus Dynamic Data Information Fusion Analysis using DDDAS for Cyber Security Trust

Erik Blasch
Youssif Al-Nashif
Salim Hariri

Originally published as Blasch E., Al-Nashif Y., Hariri S., *Static versus Dynamic Data Information Fusion analysis using DDDAS for Cyber Security Trust*, in Proc. of 14th International Conference on Computational Science (ICCS 2014), 2014, and reprinted with permission.

Abstract

Information fusion includes signals, features, and decision-level analysis over various types of data including imagery, text, and cyber security detection. With the maturity of data processing, the explosion of big data, and the need for user acceptance; the Dynamic Data-Driven Application System (DDDAS) philosophy fosters insights into the usability of information systems solutions. In this paper, we explore a notion of an adaptive adjustment of secure communication trust analysis that seeks a balance between standard static solutions versus dynamic-data driven updates. A use case is provided in determining trust for a cyber security scenario exploring comparisons of Bayesian versus evidential reasoning for dynamic security detection updates. Using the evidential reasoning proportional conflict redistribution (PCR) method, we demonstrate improved trust for dynamically changing detections of denial of service attacks.

1 Introduction

Information fusion (Blasch, *et al*., 2012) has a well-documented following of different methods, processes, and techniques emerging from control, probability, and communication theories. Information fusion systems designs require methods for big data analysis, secure communications, and support to end users. Current information fusion systems use probability, estimation, and signal processing. Extending theses techniques to operational needs requires an assessment of some of the fundamental assumptions such as secure communications over various data, applications, and systems. Specifically, the key focus of this paper is based on the question of measuring trust in static versus dynamic information fusion systems.

Static versus dynamic information fusion comes from three perspectives such as data, models, and processing. As related to information fusion techniques, many studies exist on centralized versus distributed processing, single versus multiple models, and stovepipe versus multi-modal data. In each case, static information fusion rests in centralized processing from single model estimation over a single source of data. On the other extreme is distributed processing, using multiple-models over multi-modal data; which in reality is supposed to cover the entire gamut of big data solutions captured in large-scale systems designs. In reality, with such an ambitious goal, there are always fundamental

assumptions that tailor the system design to the user needs. For example, a system could be designed to capture all image data being collected from surveillance sensors; however filtering collections over a specific area, for a designated time internal, at a given frequency helps to refine answers to user requests. Thus, as a user selects the details of importance, responses should be accessible, complete, and trustworthy.

Dynamic information fusion is a key analysis of the paper of which we focus on trust. If a machine is processing all the data, then time and usability constraints cannot be satisfied. Thus, either the user or the machine must determine the appropriate set of data, models, and processing that is needed for a specific application. Trust analysis is required to determine security and reliability constraints, and DDDAS provides a fresh look at the balance between static and dynamic information fusion. In this paper, we explore the notions of dynamic information fusion towards decision making as cyber detections change.

In Section 2 we overview information fusion and DDDAS. Section 3 discusses the notions of trust as a means to balance between information fusion and dynamic data detections. Section 4 compares Bayesian versus evidential reasoning. Section 5 provides a use-case for analysis for cyber trust and Section 6 provides conclusions..

2 Information Fusion and DDDAS

Information fusion and DDDAS overlap in many areas such as data measurements, statistical reasoning, and software development for various applications. Recently, there is an interest in both communities to address big data, software structures, and user applications. The intersection of these areas includes methods of information management (Blasch, 2006) in assessing trust in data access, dynamic processing, and distribution for applications-based end users.

2.1 Information Fusion

The *Data Fusion Information Group* (DFIG) model, shown in Figure 1, provides the various attributes of an information fusion systems design. Information fusion concepts are divided between Low-level Information Fusion (LLIF) and High-level Information Fusion (HLIF) (Blasch, *et al.*, 2012). LLIF (L0-1) composes data registration (Level 0 [L0]) and explicit object assessment (L1) such as an aircraft location and identity (Yang, 2009). HLIF (L2-6) composes much of the open discussions in the last decade. The levels, to denote processing, include situation (L2) and impact (L3) assessment with resource (L4), user (L5) (Blasch, 2002), and mission (L6) refinement (Blasch, 2005). Here we focus on Level 5 fusion by addressing cyber security trust in systems design.

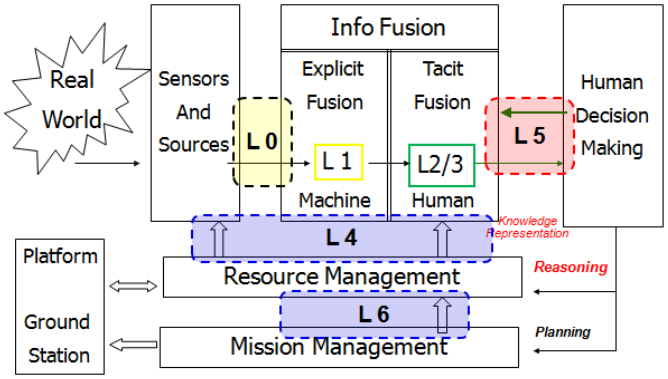


Figure 1. DFIG Information Fusion model (L = Information Fusion Level).

Data access for information fusion requires an information management (IM) model of the enterprise architecture, as shown in Figure 2. The IM model illustrates the coordination and flow of data through the enterprise with the various layers (Blasch, *et al*., 2012).

People or autonomous agents interact with the managed information enterprise environment by producing and consuming information. Various actors and their activities/services within an IM enterprise surround the IM model that transforms data into information. Within the IM model, there are various services that are needed to process the managed information objects (MIOs). Security is the first level of interaction between users and data.

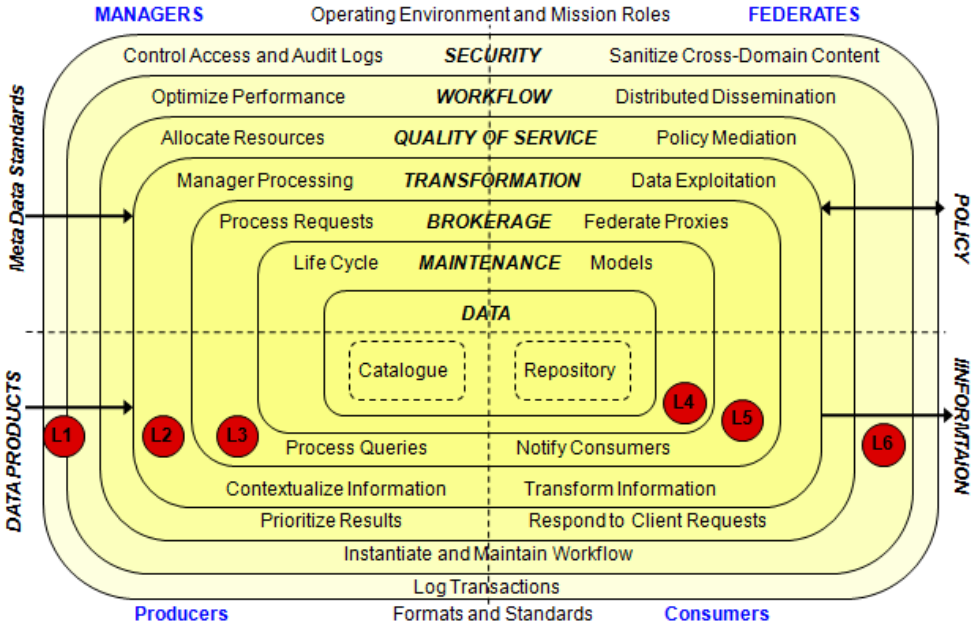


Figure 2. Information Management (IM) Model.

A set of service layers are defined that use artifacts to perform specific services. An artifact is a piece of information that is acted upon by a service or that influences the behavior of the service (e.g., a policy). The service layers defined by the model are: Security, Workflow, Quality of Service (QoS), Transformation, Brokerage, and Maintenance. These services are intelligent agents that utilize the information space within the architecture, such as cloud computing and machine analytics. Access to the data requires secure communications which is dynamic, data-type driven, and application specific.

2.2 Dynamic Data Driven Application Systems (DDDAS)

DDDAS is focused on applications modeling (scenarios), mathematical and statistical algorithms (theory), measurement systems, and systems software as shown in Figure 3. For a systems application, user mission needs drive data access over the scenarios. The available data is processed from measurements to information using theoretical principles. The data-driven results are presented to the user through visualizations; however the trust in the data is compounded by data quality, the model fidelity, and systems availability of which software is an integral part to a systems application.

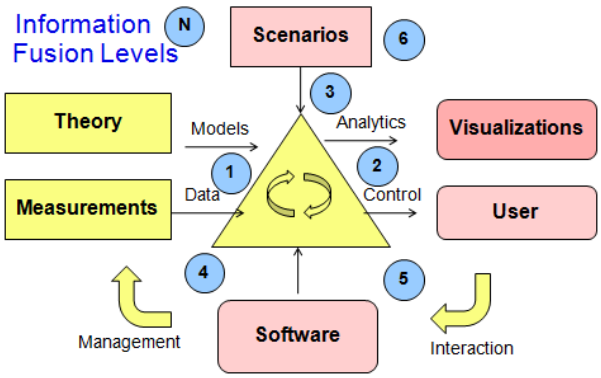


Figure 3. DDDAS Aligned with Information Fusion.

Using a cyber example for DDDAS, the application is secure data communications to meet mission needs (L6). While not a one-to-one mapping, it can be assumed that data management, driven by scenarios, identifies cyber threat attacks (L3) such as denial of service attacks. The theory and measurements come from the models of normal behavior (L1) which use computational methods to support cyber situation awareness (L2) visualization. The user (L5) interacts with the machine through data management (L4), as new measurements arrive. Current research seeks distributed, faster, and more reliable communication systems to enable such processing and coordination between the man and their machines, however, measurement of trust is paramount.

3 Trust in Information Processing

Several theories and working models of trust in automation have been proposed. Information which is presented for decision-aiding is not uniformly trusted and incorporated into situation awareness. Three proposed increasing levels, or 'stages of trust', for human-human interactions include: Predictability, Dependability, and Faith (Rempel, *et al.*, 1985). Participants progress through these stages over time in a relationship. The same was anticipated in human-automation interactions, either via training or experience. The main idea is that as trust develops, people will make decisions based upon the trust that the system will continue to behave in new situations as it has demonstrated in the past. Building upon Rempel's stages, (Muir & Moray, 1996) postulated that

Trust = Predictability + Dependability + Faith + Competence + Responsibility + Reliability

and further defined the construct of *Distrust***:** which (1) can be caused by operator feeling that the automation is undependable, unreliable, unpredictable, etc. and a (2) set of dimensions related to automation failures, which may cause distrust in automated systems (location of failure, causes of failure or corruption, time patterns of failure).

Table 1, adapted below from (Muir & Moray, 1996), depicts the quadrant of trust and distrust behaviors with respect to good or poor quality of the automation. Basically, the outcome of a wrong decision to trust the automation is worse than the outcome of a wrong decision to not trust the automation. Hence, security is enforced to not trust a poor decision.

Operator's trust & allocation of function	Quality of the automation **'Good'**	**'Poor'**
Trusts and uses the automation	*Appropriate Trust* (optimize system performance)	*False Trust* (risk automated disaster)
Distrusts and rejects the automation	*False Distrust* (lose benefits of automation, inc. workload)	*Appropriate Distrust* (optimize system performance)

Table 1: Trust, Distrust, and Mistrust, (adapted from Muir and Moray, 1996)

Trust in the automation clearly impacts a user mental model of secure communications. Therefore, dynamic models must be devised to account for different levels of attention, trust, and interactions in Human in the Loop (HIL) and Human on the Loop (HOL) designs. A user must be given permission to refine the assessment for final decision for validity and reliability of the information presented. User Trust issues then are confidence (correct detection), security (impacts), integrity (what you know), dependability (timely), reliable (accurate), controllability, familiar (practice and training), and consistent (reliable).

Trust in information processing involves many issues; however, here we focus on the development of a cyber domain trust stack as shown in Figure 4. The trust stack composes policies, trust authority, collecting raw metrics and behavior analysis, leading to authentication and authorization, and then secure communications. Similar to the information management model, polices are important to determine whether data access is available. Likewise, sensor management gets access to raw metrics (Blasch, 2004) that need to be analyzed for situation awareness. The problem not being full addressed is the impeding results for secure communications. In what follows, we discuss the main functions to be provided by each layer in the trust stack shown in Figure 4.

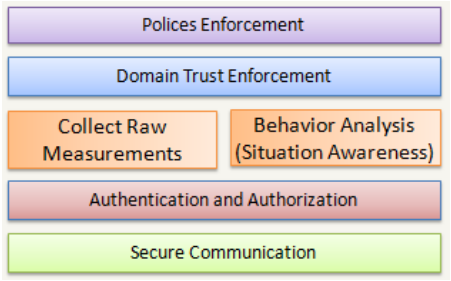


Figure 4. Trust Stack.

3.1 Secure Communications, Authentication, and Authorization

Secure communications is an important property to guarantee the confidentiality and integrity of the messages used to evaluate trust in the system. Certificates are used to verify the identify of communicating end-devices (Kaliski, 1993). The communication channel is encrypted using DES (Data Encryption Standard, 2010) in CFB64 (Cipher Feedback) mode. In this CFB mode, the first 8 bytes of the key generated used to encrypt the first block of data. This encrypted data is then used as a key for the second block. This process is repeated until the last block is encrypted. The DES is still used in legacy virtual private networks (VPNs) and could benefit from a DDDAS trust analysis even used with multiple protocol authentication systems such as Kerberos.

Multiple protocols have been developed over the years for password-based authentication, biometric authentication, and remote user authentication. In order to evaluate the trust of different entities with many users, multiple systems, and multiple domains, we assume the use of remote user authentication. Remote Authentication Dial-In User Service (RADIUS) (Willens, *et al*, 2000) is a

famous client/server protocol to allow remote entities to communicate with a server to authenticate remote users. RADIUS gives organization ability to maintain user profiles in a specific database that the remote servers share.

The Domain Trust Enforcement (DTE) agent performs the authorization process for the end-to-end adaptive trust. Based on the results of the authentication process and the received trust level, the DTE agent grants or denies authorization to access the resources, i.e., allow or deny the communication between the different entities.

3.2 Collecting Raw Measurements

Much software, both commercial and open source, are available and provide important health and security information, such as Nagios (Nass, 2009). This information can be used to extract metrics that can be used to evaluate the trust of different entities. These metrics can be divided into multiple categories based on their source: User, Application, Machine, Connection, or Security Software Alerts. In order to evaluate the trust, the metrics need to be quantified and normalized (e.g., between 0 and 1) to a common scale. Table 2 shows a set of measured metrics and their quantification function and Figure 5 shows these categories with some example metrics.

Category	***Metric***	***Quantification***
User	Password Strength	$\begin{cases} 0, & \text{Password Length} < 8 \\ 0.1 + 0.9 \cdot \dfrac{\text{Password Length}}{\text{Maximum Password Length}}, & \text{Otherwise} \end{cases}$
User	Days since last password change	$\begin{cases} 0, & \#\text{days} > \text{Maximum Number Of Days} \\ 1 - \dfrac{\#\text{days}}{\text{Maximum Number of Days}}, & \text{Otherwise} \end{cases}$
User	Number of authentication failures	$\begin{cases} 0, & \#\text{failures} > \text{Maximum Number Of Allowed Failures} \\ 1 - \dfrac{\#\text{failures}}{\text{Maximum Number Of Allowed Failures}}, & \text{Otherwise} \end{cases}$
User	Lock Outs	$\begin{cases} 0, & \#\text{Lock Outs} > \text{Maximum Number Of Allowed Lock Outs} \\ 1 - \dfrac{\#\text{Lock Outs}}{\text{Maximum Number Of Allowed Lock Outs}}, & \text{Otherwise} \end{cases}$
Application	Developer Reputation	$\dfrac{\text{Reputation}}{\text{Maximum Reputation}}$
Application	Who manages the software	$\begin{cases} 1, & \text{Global Adminstrator} \\ 0.5, & \text{Local Administrator} \\ 0, & \text{No Administrator} \end{cases}$
Connection	Number of hops	$\begin{cases} 0, & \#\text{Hops} > \text{Maximum Number Of Hops} \\ 1 - \dfrac{\#\text{Hops}}{\text{Maximum Number of Hops}}, & \text{Otherwise} \end{cases}$
Connection	Number of discarded Packets	$\begin{cases} 0, & \#\text{Discarded Packet} > \text{Maximum \#Discarded Packet} \\ 1 - \dfrac{\#\text{Discarded Packet}}{\text{Maximum \#Discarded Packet}}, & \text{Otherwise} \end{cases}$
Machine	Firmware version	$\begin{cases} 1, & \text{Up to date} \\ 0.5, & \text{1 Version Behind} \\ 0, & \text{Otherwise} \end{cases}$
Machine	Shared Folders	$\begin{cases} 1, & \text{No Shared Folders} \\ 0.5, & \text{Shared User Folders} \\ 0, & \text{Shared System Folders} \end{cases}$
Analyzer	Integrity Check	$\begin{cases} 1, & \text{No Probelm} \\ 0.5, & \text{Problem in user data} \\ 0, & \text{Problem in system integrity} \end{cases}$
Analyzer	Virus Alerts	$\begin{cases} 1, & \text{No Alert} \\ 0.5, & \text{Virus Found in a document} \\ 0.25, & \text{Virus Found in an executable} \\ 0, & \text{worm found} \end{cases}$

Table 2: Examples of metric quantification

3.3 Behavior Analysis

Behavior analysis techniques apply statistical and data mining techniques to determine the current operating zone of the execution environment (situation awareness) and also project its behavior in the near future. The operating point (OP) of an environment can be defined as a point in an *n*-dimensional space with respect to well-defined attributes. An acceptable operating zone can be defined by combining the normal operating values for each attribute. At runtime, the operating point moves from one zone to another and that point might move to a zone where the environment does not meet its trust and security requirements. We use these movements in the OP to adjust the trust value of the current environment as will be discussed in further detail in the Domain Trust Authority section. By continuously performing behavior analysis of the environment, we can then proactively predict and detect the anomalous behaviors that might have been caused by malicious attacks. Furthermore, once it is determined that the environment's operating point is moving outside the normal zone, it will adopt its trust value and then determine the appropriate proactive management techniques that can bring back the environment situation to a normal operating zone.

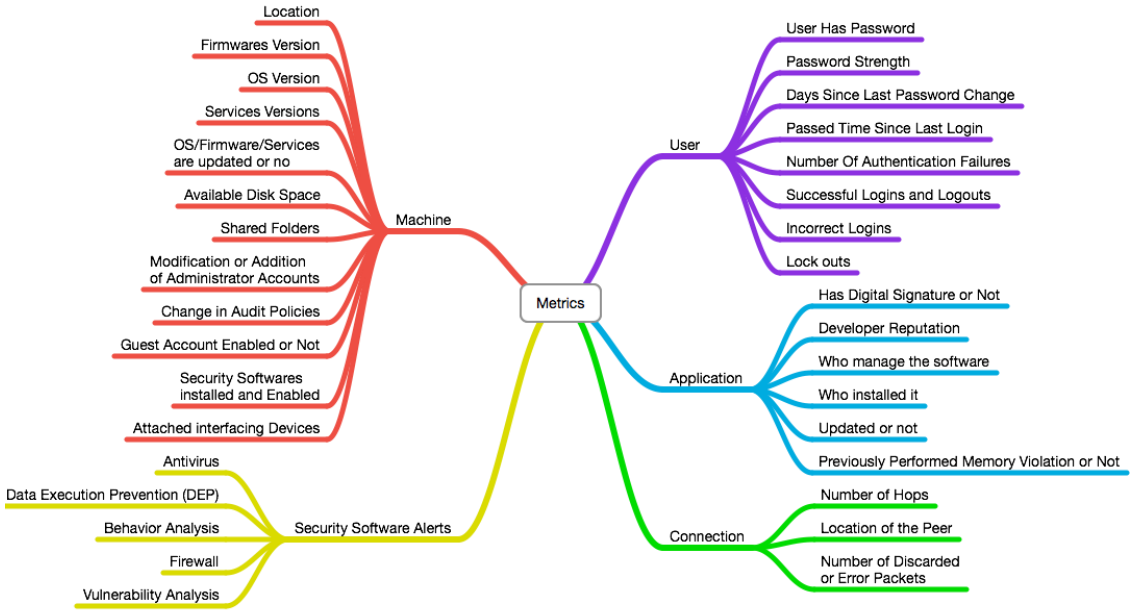


Figure 5. Trust Metrics.

3.4 Domain Trust Authority

DTA evaluates the end-to-end trust over secure communications. It defines a tuple (machine, application, user, data) to be an entity and all communications among entities has a certain context. Thus authentication is conducted per entity. Every entity has a trust level associated with it. In order to measure the trust, trust's metrics are introduced, and they take values between 0 and 1. Where 0 represents the distrust and 1 represent the blind or full trust. The trust measurements for all entities are stored in an entity call Trust Authority. The NIST standard SP 800-53 (NIST, 2010) is used and it defines four levels of trust:

Level	Distrust	Low Trust	Moderate	High Trust
Trust Value	0.00	0.33	0.66	1.00

Initially, a risk and impact analysis is performed to quantify the impact of each component on the overall operations of the network. Common Vulnerabilities and Exposures (CVE) and Common Vulnerability Scoring System (CVSS) are used to evaluate the initial impact for both software and the environment, and reputations of the users are used to assign their initial impacts. Based on the initial impact analysis, the initial trust values for each entity is determined. The risk and impact analysis performed is in consistence with the NIST "Recommended Security Controls for Federal Information Systems and Organizations" report. According to the NIST report, risk measures the extent to which entities are threatened by circumstances or events. The risk is a function of impact and its probability of occurrence. Risks arise from the loss of confidentiality, integrity, and/or availability of information and resources. Thus the initial trust T can be viewed as an inverse function of the risk R:

$$T = 1 / R \tag{1}$$

Where the risk of an entity i is a function of the impact imp:

$$R_i = imp_i(\text{confidentiality}) \bullet Pr\, imp_i(\text{confidentiality}) + imp_i(\text{integrity}) \bullet Pr\, imp_i(\text{integrity}) + imp_i(\text{availability}) \bullet Pr\, imp_i(\text{availability}) \tag{2}$$

When a new entity is added, it has to register with the Mutual Authentication (MA) module and then its initial trust value can be quantified according to Equations 1 and 2.

Verify Trust

When an entity communicates with another entity, an Autonomic Trust Management (ATM) agent obtains the trust level of the entity that needs to interact with from the Trust Authority (TA), see Figure 6. If the trust level of the remote entity is below the minimum required trust level set in the policies, then the communication is dropped. By continuously checking with TA module, any interacting entities will not be able to communicate if they do not meet the end-to-end trust policies. Once the component trust level is verified, they can proceed and interact securely using the secure communications.

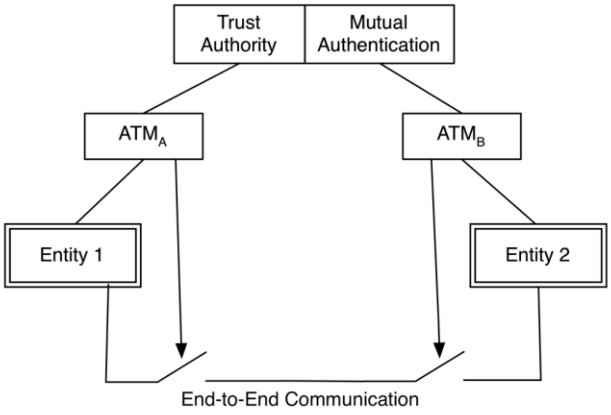


Figure 6. Adaptive End-to-End Trust

Adaptive Trust

The trust value assigned to each component is not static and is updated continuously. The Trust Authority module is the one responsible for re-evaluating the trust at runtime. As mentioned in the previous section, the trust is measured per entity and the trust levels are between 0 and 1.

$$T(E) \in [0, 1] \tag{3}$$

Each interaction between entities is governed by a context *C*. Thus, trust level for entities is computed per context:

$$T(E, C) \in [0, 1] \tag{4}$$

A Forgiveness Factor, *F*, is assigned to provide an adaptive mechanism for compromised entities to start gaining trust after all existing vulnerabilities have been fixed. Based on the impact of the entity on the overall operations, we can control the time it takes for that entity to recover its trust level. Monitoring, measuring, and quantifying trust metrics are required, and they are performed by the ATM. M_i will denote the collected trust metric, where *i* is the metric identifier. The function $m_i()$ is a quantifying function that returns a measurement between 0 and 1 for the metric M_i.

The overall trust for an entity is computed using two types of trust: 1) self-measured trust and 2) reputation-measured trust. The self-measured trust T_s is the trust that is evaluated based on the measurement performed by the ATM agent that manages the entity. While the reputation-measured trust, T_p is based on the trust metrics collected from peers based on a previous recent interaction with the entity for which the trust is being re-evaluated. The T_s and T_p are given by following equations:

$$\begin{aligned} T_{\mathrm{S}}(E, C) &= T(ATM_{\mathrm{E}}, C) \cdot \sum_{i=1}^{\mathrm{L}} I_{\mathrm{i}}(C) \cdot m_{\mathrm{i}}(M_{\mathrm{i}}) \\ T_{\mathrm{P}}(E, C) &= \frac{1}{K} \sum_{j=1}^{\mathrm{K}} T(ATM_{\mathrm{j}}, C) \cdot \sum_{i=1}^{\mathrm{L}} I_{\mathrm{i}}(C) \cdot m_{\mathrm{i}}(M_{\mathrm{i}}) \end{aligned} \tag{5}$$

The values of the metric weight I_i for metric *i* is determined based on the feature selection technique, where:

$$\sum_{i=1}^{\mathrm{L}} I_{\mathrm{i}}(C) = 1 \tag{6}$$

Based on the context and the type of operations, the end-to-end trust is evaluated using three trust evaluation strategies: Optimistic, Pessimistic, and Average. The end-to-end trust for each strategy can be evaluated as follows:

Trust Confidence	**Trust Evaluation Strategy**
Optimistic	$T(E, C) = max\ \{T_{\mathrm{S}}(E, C), T_{\mathrm{P}}(E, C)\}$
Average	$T(E, C) = ave\ \{T_{\mathrm{S}}(E, C), T_{\mathrm{P}}(E, C)\}$
Pessimistic	$T(E, C) = min\ \{T_{\mathrm{S}}(E, C), T_{\mathrm{P}}(E, C)\}$

Once $T(E,C)$ is computed, then it is mapped to the nearest of trust level: (High, Moderate, Low, and None).

The Trust Authority module continuously evaluates the trust for all components and their entities whenever new metrics are obtained from the ATM agents that require an update to entity trust evaluation above depending on the trust evaluation strategy. Various reasoning evaluation strategies exist, such as that of Bayesian, Evidential Reasoning, and Belief Functions (Blasch, *et al*, 2013), that can be used to evaluate trust.

In a DDDAS cyber environment, there are many levels of information fusion, but to build a trustworthy DDDAS environment, we need to check the trust of each level of information fusion. The Domain Trust Authority is the place to verify the trust of each entity passing information within the DDDAS environment. When the trust level drops below certain threshold; the incoming data can be

dropped to enable secure communications. What follows are the DDDAS theory, simulations, measurements, and software analysis for Information fusion levels of cyber data, situation/behavior assessment, information management, and user refinement.

3.5 Bayes versus Evidential Reasoning

A fundamental technique for data fusion is Bayes Rule. Recently, (Dezert, *et al*., 2012) has shown that Dempster's rule is consistent with probability calculus and Bayesian reasoning if and only if the prior $P(X)$ is uniform. However, when the $P(X)$ is not uniform, then Dempster's rule gives a different result. Both (Yen, 1986) and (Mahler, 1996) developed methods to account for non-uniform priors. Others have also tried to compare Bayes and evidential reasoning (ER) methods (Mahler, 2005, Blasch, *et al*., 2013). Assuming that we have multiple measurements $Z = \{Z_1, Z_2, \ldots, Z_N\}$ for cyber detection D being monitored, Bayesian and ER methods are developed next.

3.6 Relating Bayes to Evidential Reasoning

Assuming conditional independence, one has the Bayes method:

$$P(X \mid Z_1 \cap Z_2) = \frac{P(X \mid Z_1)\, P(X \mid Z_2) / P(X)}{\sum_{i=1}^{N} P(X_i \mid Z_1)\, P(X_i \mid Z_2) / P(X_i)} \tag{7}$$

With no information from Z_1 or Z_2, then $P(X \mid Z_1, Z_2) = P(X)$. Without Z_2, then $P(X \mid Z_1, Z_2) = P(X \mid Z_1)$ and without Z_1, then $P(X \mid Z_1, Z_2) = P(X \mid Z_2)$. Using Dezert's formulation, then the denominator can be expressed as a normalization coefficient:

$$m_{12}(\varnothing) = 1 - \sum_{X_i ; X_j \mid X_i \cap X_j} P(X_i \mid Z_1)\, P(X_i \mid Z_2) \tag{8}$$

Using this relation, then the total probability mass of the conflicting information is

$$P(X \mid Z_1 \cap Z_2) = \frac{1}{1 - m_{12}(\varnothing)} \bullet P(X \mid Z_1)\, P(X \mid Z_2) \tag{9}$$

which corresponds to Dempster's rule of combination using Bayesian belief masses with uniform priors. When the prior's are not uniform, then Dempster's rule is not consistent with Bayes' Rule. For example, let $m_0(X) = P(X)$, $m_1(X) = P(X \mid Z_1)$, and $m_2(X) = P(X \mid Z_2)$, then

$$m(X) = \frac{m_0(X)\; m_1(X)\; m_2(X)}{1 - m_{012}(\varnothing)} = \frac{P(X)\;\; P(X \mid Z_1)\;\; P(X \mid Z_2)}{\sum_{i=1}^{N} P(X_i)\;\; P(X_i \mid Z_1)\;\; P(X_i \mid Z_2)} \tag{10}$$

Thus, methods are needed to deal with non-uniform priors and appropriately redistribute the conflicting masses.

3.7 Proportional Conflict Redistribution

Recent advances in DS methods include *Dezert-Smarandache Theory* (DSmT). DSmT is an extension to the Dempster-Shafer method of evidential reasoning which has been detailed in numerous papers and texts: *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3 (Dezert, *et al*., 2009). In (Dezert, *et al*., 2002) introduced the methods for the

reasoning and in presented the hyper power-set notation for DSmT (Dezert, *et al*., 2003). Recent applications include the DSmT Proportional Conflict Redistribution rule 5 (PCR5) applied to target tracking (Blasch, 2013).

The key contributions of DSmT are the redistributions of masses such that no refinement of the frame Θ is possible unless a series of constraints are known. For example, Shafer's model (Shafer, 1976) is the most constrained DSm hybrid model in DSmT. Since Shafer's model, authors have continued to refine the method to more precisely address the combination of conflicting beliefs (Josang, *et al*., 2006) and generalization of the combination rules (Smaradache, *et al*., 2005, Daniel, 2006). An adaptive combination rule (Florea, *et al*., 2006) and rules for quantitative and qualitative combinations (Martin, 2008) have been proposed. Recent examples for sensor applications include electronic support measures, (Djiknavorian, *et al*., 2010), physiological monitoring sensors (Lee, *et al*., 2010), and seismic-acoustic sensing (Blasch, *et al*., 2011).

Here we use the *Proportional Conflict Redistribution* rule no. 5 (PCR5)[*]. We replace Smets' rule (Smets, 2005) by the more effective PCR5 to cyber detection probabilities. All details, justifications with examples on PCR*n* fusion rules and DSm transformations can be found in the DSmT compiled texts (Dezert, *et al*., 2009 Vols. 2 & 3). A comparison of the methods is shown in Figure 7.

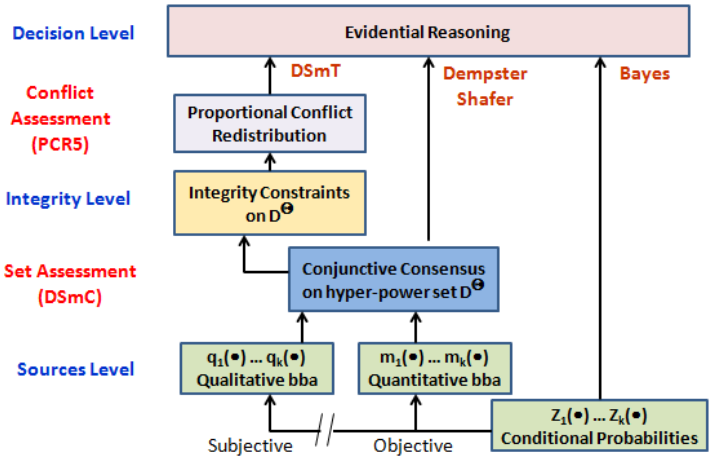


Figure 7. Comparison of Bayesian, Dempster-Shafer, and PCR5 Fusion Theories

In the DSmT framework, the PCR5 is used generally to combine the basic belief assignment (bba)'s. PCR5 transfers the conflicting mass only to the elements involved in the conflict and proportionally to their individual masses, so that the specificity of the information is entirely preserved in this fusion process. Let $m_1(.)$ and $m_2(.)$ be two independent bba's, then the PCR5 rule is defined as follows (see Dezert, *et al*., 2009, Vol. 2 for full justification and examples): $m_{PCR5}(\varnothing) = 0$ and $\forall X \in 2^\Theta \setminus \{\varnothing\}$, where $\varnothing$ is the null set and 2^Θ is the power set:

$$m_{PCR5}(X) = \sum_{\substack{X_1;\, X_2 \in 2^\Theta \\ X_1 \cap X_2 = X}} m_1(X_1) + m_2(X_2) + \sum_{\substack{X_2 \in 2^\Theta \\ X_2 \cap X = \varnothing}} \left[\frac{m_1(X_1)^2\, m_2(X_2)}{m_1(X_1) + m_2(X_2)} + \frac{m_1(X_1)\, m_2(X_2)^2}{m_1(X_1) + m_2(X_2)} \right] \quad (11)$$

[*] Note: PCR used here is from information fusion technology and not the a Platform Configuration Register (PCR) of the Trusted Platform Module (TPM) hardware technology.

where ⋂ is the interesting and all denominators in the equation above are different from zero. If a denominator is zero, that fraction is discarded. Additional properties and extensions of PCR5 for combining qualitative bba's can be found in (Dezert, 2009, Vol. 2 & 3) with examples and results. All propositions/sets are in a canonical form.

3.8 Example of DDDAS Cyber Trust Analysis

In this example, we assume that policies are accepted and that the trust stack must determine whether the dynamic data is trustworthy. The application system collects raw measurements on the data intrusion (such as denial of service attacks) and situation awareness is needed. Conventional information fusion processing would include Bayesian analysis to determine the state of the attack. However, here we use the PCR5 rule which distributes the conflicting information over the partial states. Figure 8 shows the results for a normal system being attacked and the different methods (Bayes, DS, and PCR5) to access the dynamic attack. Trust is then determined with percent improvement in analysis. Since the cyber classification of attack versus no attack is not consistent, there is some conflict in the processing of the measurement data going from an measurements of attack and vice versa. The constant changing of measurements requires acknowledgment of the change and data conflict as measured using the PCR5 method.

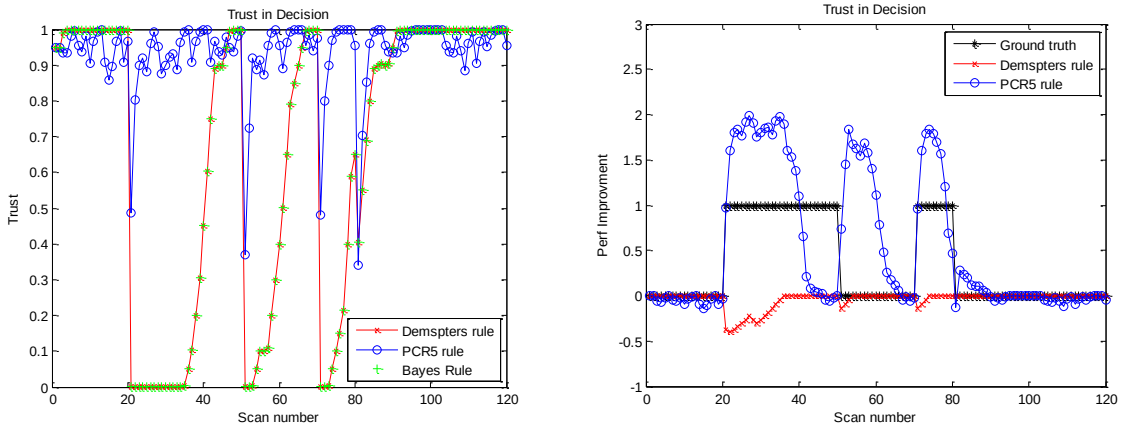


Figure 8. Results of Bayesian, Dempster-Shafer, and PCR5 Fusion Theories for trust.

The improvement of PCR5 over Bayes is shown in Figure 8 and compared with the modest improvement from DS. The average performance improvement of PCR5 is 46% and DS is 2%, which is data and application dependent. When comparing the results, it can be seen that when a system goes from a normal to an attack state, PCR5 responds quicker in analyzing the attack, resulting in maintaining trust in the decision. Such issues of data reliability, statistical credibility, and application survivability all contribute to the presentation of information to an application-based user. While the analysis is based on behavioral situation awareness, it is understood that polices and secure communications can leverage this information for domain trust analysis and authentication and authorization that can map measurements to software requirements.

3.9 Policies Enforcement

Policies are an important component of cyber trust (Blasch, 2012) as shown in Figure 9. As an example, a policy is administered for retrieval of information. Policy information determines the attributes for decisions. Determining the decision leads to enforcement. Such a decision is based on trust processing from which effective enforcement can support secure communications.

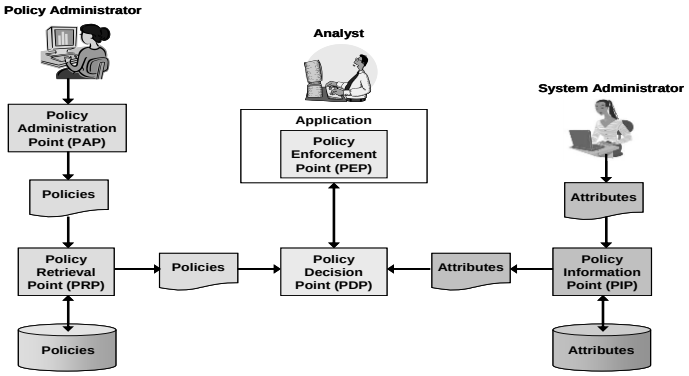


Figure 9. Policy-Based Fusion of Information requiring Trust (Blasch, 2012)

There are many possible information fusion strategies to enable data access from policies. Here we demonstrate an analysis of Bayesian versus evidential reasoning for determining cyber situation awareness trust. Future work includes threat intent (Shen, *et al*., 2009), impact assessment (Shen, *et al*., 2007), transition behaviors (Du, *et al*., 2011) and developing advanced forensics analysis (Yu, *et al*., 2013).

4 Conclusions

Information fusion (IF) and Dynamic Data-Driven Application Systems (DDDAS) are emerging techniques to deal with big data, multiple models, and decision making. One topic of interest to both fields of study is a measure of trust. In this paper, we explored a system for cyber security fusion which addresses system-level application issues of model building, data analysis, and polices for application trust. IF and data-driven applications utilize a common framework of probability analysis and here we explored a novel technique of PCR5 that builds on Bayesian and Dempster-Shafer theory to determine trust. Future research would include real world data, complete analysis of the trust stack, and sensitivity of models/measurements in secure cyber situation awareness trust analysis.

Acknowledgements

This work is partially supported by AFOSR DDDAS award number FA95550-12-1-0241, and National Science Foundation research projects NSF IIP-0758579, NCS-0855087 and IIP-1127873.

References

Blasch, E., Plano, S. (2002) "JDL Level 5 Fusion model 'user refinement' issues and applications in group Tracking," *Proc. SPIE*, Vol. 4729.

Blasch, E., Plano, S. (2003) "Level 5: User Refinement to aid the Fusion Process," *Proc. of SPIE,* 5099, 2003.

Blasch, E., Pribilski, M., Daughtery, B., Roscoe, B., and Gunsett, J. (2004) "Fusion Metrics for Dynamic Situation Analysis," *Proc. of SPIE*, Vol. 5429.

Blasch, E., Plano, S. (2005) "DFIG Level 5 (User Refinement) issues supporting Situational Assessment Reasoning," *Int. Conf. on Info Fusion*.

Blasch, E. (2006) "Level 5 (User Refinement) issues supporting Information Fusion Management," *Int. Conf. on Info Fusion*.

Blasch, E., Kadar, I., Salerno, J., Kokar, M. M., Das, S., Powell, *et al*.. (2006) "Issues and Challenges in Situation Assessment (Level 2 Fusion)," *J. of Advances in Information Fusion,* Vol. 1, No. 2, pp. 122 - 139, Dec.

Blasch, E., Dezert, J., Valin, P. (2011) "DSMT Applied to Seismic and Acoustic Sensor Fusion," *Proc. IEEE Nat. Aerospace Electronics Conf (NAECON).*

Blasch, E., Bosse, E., Lambert, D. A. (2012), *High-Level Information Fusion Management and Systems Design*, Artech House, Norwood, MA.

Blasch, E., Dezert, J., Pannetier, B. (2013) "Overview of Dempster-Shafer and Belief Function Tracking Methods," *Proc. SPIE*, Vol. 8745,

Blasch, E., Steinberg, A., Das, S., Llinas, J., Chong, C.-Y., Kessler, O., Waltz, E., White, F., (2013) "Revisiting the JDL model for information Exploitation," *Int'l Conf. on Info Fusion*.

Blasch, E. (2013) "Enhanced Air Operations Using JView for an Air-Ground Fused Situation Awareness UDOP," *AIAA/IEEE Digital Avionics Systems Conference*, Oct..

Chen, G., Shen, D., Kwan, C., Cruz, J., *et al*., (2007) "Game Theoretic Approach to Threat Prediction and Situation Awareness," *Journal of Advances in Information Fusion*, Vol. 2, No. 1, 1-14, June.

Culbertson, J., and Sturtz, K., (2013) "A Categorical Foundation for Bayesian Probability," *Applied Categorical Structures*.

Daniel, M., (2006) "Generalization of the Classic Combination Rules to DSm Hyper-Power Sets," *Information & Security, An Int'l J.*, Vol. 20.

Data Encryption Standard (2010), http://blog.fpmurphy.com/2010/04/openssl-des-api.html

Dezert, J. (2002) "Foundations for a new theory of plausible and paradoxical reasoning," *Information & Security, An Int'l J.*, ed. by Prof. Tzv. Semerdjiev, Vol. 9.

Dezert, J. Smarandache, F. **(**2003) "On the generation of hyper-powersets for the DSmT," *Int. Conf. on Info Fusion*.

Dezert, J. Smarandache, F., (2009) *Advances and applications of DSmT for information fusion (Collected works)*, Vols. 1-3, American Research Press, http://www.gallup.unm.edu/~smarandache/DSmT.htm

Dezert, J. (2012) "Non-Bayesian Reasoning for Information Fusion – A Tribute to Lofti Zadeh," *submitted to J. of Adv. of Information Fusion*.

Djiknavorian, P., Grenier, D., Valin, P. (2010) "Approximation in DSm theory for fusing ESM reports," *Int. Workshop on Belief functions* 2010, April.

Du, H. Yang, S. J. (2011) "Characterizing Transition Behaviors in Internet Attack Sequences," in *IEEE ICCCN'11*.

Dsouza, G., Rodriguez, G., Al-Nashif, Y., Hariri, S. (2013) "Resilient Dynamic Data Driven Application Systems (rDDDAS)," *International Conference on Computational Science*.

Dsouza, G., Hariri, S., Al-Nashif, Y., Rodriguez, G. (2013) "Building resilient cloud services using DDDAS and moving target defense," *Int. J. Cloud Computing*..

Florea, M. C., Dezert, J., Valin, P., Smarandache, F., Jousselme, A-L., (2006) "Adaptive combination rule and proportional conflict redistribution rule for information fusion," *COGIS '06 Conf.*,

Josang, A., Daniel, M. (2006) "Strategies for Combining Conflict Dogmatic Beliefs," *Int. Conf. on Info Fusion*.

Kaliski, B. (1993) "A Survey of Encryption Standards," *IEEE Micro*, Issue, 6, December.

Lee, Z. H., Choir, J. S., Elmasri, R. (2010). "A Static Evidential Network for Context Reasoning in Home-Based Care," *IEEE Trans. Sys., Man, and Cyber-Part A; Sys & Humans*, Vol. 40, No. 6, Nov.

Mahler, R.P. (1996) "Combining ambiguous evidence with respect to ambiguous a priori knowledge, I: Boolean logic," *IEEE Trans. Sys., Man & Cyber., Part A*, Vol. 26, pp. 27–41.

Mahler, R., (2005) "Can the Bayesian and Dempster-Shafer approaches be reconciled? Yes," *Int'l Conf. on Information Fusion*.

Martin, A., Osswald, C., Dezert, J., Smarandache, F. (2008) "General Combination Rules for Qualitative and Quantitative Beliefs," *J. of Advances in Information Fusion*, Vol. 3, No. 2, Dec.

Muir, B. and Moray, N. (1996) "Trust in automation: Part II. Experimental studies of trust and human Intervention in a process control simulation," *Ergonomics*, 39 (3), 429-460.

Nass, S. J., Levit, L. A., Gostin, L. O. (2009). Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research, National Academies Press.

NIST, (revision, 2010) "Recommended Security Controls for Federal Information Systems and Organizations," NIST Special Publication 800-53, Revision 3.

Rempel, J. K., Holmes, J. G., and Zanna, M. P. (1985) "Trust in Close Relationships," *Journal of Personality and Social Psychology*, 49 (1), 95-112.

Shafer, G. (1976) *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton Univ. Press.

Shen, D., Chen, G. et al., (2007) "Strategies Comparison for Game Theoretic Cyber Situational Awareness and Impact Assessment," *Int. Conf. on Info Fusion*..

Shen, D., Chen, G., *et al*. (2009) "An Adaptive Markov Game Model for Cyber Threat Intent Inference", invited Ch. 21 in *Theory and Novel Applications of Machine Learning*, M. J. Er and Y. Zhou. (Eds.), IN-TECH.

Smaradache, F., Dezert, J., (2005) "Information fusion based on new proportional conflict redistribution rules," *Int. Conf. Inf. Fusion*.

Smets, P., (2005) "Analyzing the Combination of Conflicting Belief Functions," *Int. Conf. on Info Fusion*.

Willens, S., *et al*, (2000) Remote Authentication Dial-In User Service (RADIUS), accessed at http://tools.ietf.org/search/rfc2865

Yang, C., Blasch, E. (2009) "Kalman Filtering with Nonlinear State Constraints," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 45, No. 1, 70-84, Jan.

Yen, J. (1986) "A reasoning model based on the extended Dempster Shafer theory," *Nat Conf. on Artificial Intelligence*.

Yu, W., Fu, X., *et al*. (2013) "On Effectiveness of Hopping-Based Techniques for Network Forensic Traceback," *Int'l J. of Networked and Distributed Computing*, Vol. 1, No. 3, 2013.

Part 3:

List of References related with DSmT

Please visit regularly the following web pages to stay tuned with advances and applications of DSmT:

http://fs.gallup.unm.edu//DSmT.htm

http://www.onera.fr/staff/jean-dezert

2014 Publications

1. 2014 - Performance of M-ary soft fusion systems using simulated human responses, D.J. Bucci, S. Acharya, M. Kam, Proc. of 17th Int Conf on Information Fusion, Salamanca, Spain, July 7-10, 2014.
2. 2014 - Han D., Dezert J., Yang Y., New Distance Measures of Evidence based on Belief Intervals, in Proc. of Belief 2014 Conf. Oxford, UK, Sept. 26-29, 2014.
3. 2014 - Static versus Dynamic Data Information Fusion analysis using DDDAS for Cyber Security Trust, E. Blasch, Y. Al-Nashif, S. Hariri, In Proc. of International Conference on Computational Science (ICCS 2014), Procedia Computer Science, Vol. 29, pp. 1299–1313, 2014.
4. 2014 - Decision-Making Level Fusion Based on DSmT for Multi-Sensor Life Detection Platform, Jingsong Yang, Xun Chen, Sensors & Transducers, Vol. 164, Issue 2, pp. 31-35, February 2014.
5. 2014 - Smarandache F., Dezert J., Martin A., Comments on the paper "An Alternative Combination Rule for Evidential Reasoning" by Sebbak et al., published in Fusion 2014 Conference, Salamanca, Spain, July 2014, Bulletin of Pure and Applied Sciences, Volume 33 E (Math & Stat.) Issue (No.2) 2014, pp. 91-94.
6. 2014 - Tchamova A., Dezert J., Performance evaluation of fuzzy-based fusion rules for tracking applications, in International Journal of Reasoning-based Intelligent Systems (IJRIS), Vol. 6, No. 3/4, pp. 126--135, 2014.
7. 2014 - Contextual reliability discounting in welding process diagnostic based on DSmT, Wojciech Jamrozik, to appear in Expert Systems, 2014 (Wiley Publishing Ltd.)
8. 2014 - New method for multiple cues fusion combined DST and DSmT, Ji-Wen Tan, Hong Zhan, Yan Wen, Wei-Xia Zhan, Information Technology Journal 13 (2), pp. 393--396, 2014.
9. 2014 - Trust metrics in information fusion, Erik Blasch, Proc. SPIE 9119, Machine Intelligence and Bio-inspired Computation: Theory and Applications VIII, 91190L Baltimore, Maryland, USA, May 5th, 2014.
10. 2014 - Uncertainty evaluation for a Dezert–Smarandache theory-based localization problem, M. Khodabandeha, A.M. Shahria, International Journal of General Systems, Volume 43, Issue 6, pp. 610-632, March 2014.
11. 2014 - Comparison of identity fusion algorithms using estimations of confusion matrices, G. Golino, A. Graziano, A. Farina, W. Mellano, F. Ciaramaglia, Proc. of 17th Int Conf on Information Fusion, Salamanca, Spain, July 7-10, 2014.
12. 2014 - A new self-adaptive fusion algorithm based on DST and DSmT, X.H. Yu, Q.-J. Zhou, Y.-L. Li, J. An, Z.-C. Liu, Proc. of 17th Int Conf on Information Fusion, Salamanca, Spain, July 7-10, 2014.
13. 2014 - Improved Method Based on DSmT and Its Applications in C4ISR System, D. Qin, Z. Miao, Y. Wang, Journal of University of Electronic Science and Technology of China, Vol. 43, No. 4, July 2014 (in Chinese).
14. 2014 - Generalized PCR combination rules under framework of DSmT, Chen Jinguang, Zhang Fen, in Application Research of Computers, Vol. 31, No. 8, pp.2346--2349, 2014.
15. 2014 - New method for multiple cues fusion combined DST with DSmT, by Ji-Wen Tan; Hong Zhan; Yan Wen; et al., Information Technology Journal , Vol. 13, No. 2, pp. 393-396, 2014.
16. 2014 - Representation of random sets for PCR rules under framework of DSmT, by Ma Lili, Zhang Fen & Chen Jinguang, Computer Engineering and Design, Vol. 35, No. 3, pp. 1046-1050, 2014.

17. 2014 - Improvement of the walking robot dynamic stability using the DSmT and the neutrosophic logic, by Vladareanu, L.; Gal, A.; Hongnian Yu; et al., in 2014 International Conference on Advanced Mechatronic Systems (ICAMechS), 10-12 Aug. 2014, Kumamoto, Japan.

2013 Publications

1. 2013 - Simultaneous fault diagnosis of the reactor coolant system based on DSm evidence theory, Jinshen Ren, Botao Jiang, Fuyu Zhao, in Proc. of 21st International Conference on Nuclear Engineering, Chengdu, China, July 29-Aug 2, 2013 (Paper No. ICONE21-16209).
2. 2013 - Lee H., A Dynamic Evidential Fusion Network for Decision Making in Multi- Robot System, International Journal of Information Processing and Management (IJIPM), Vol. 4, No. 3, pp. 156--170, 2013.
3. 2013 - J. Daniel, J.-P. Lauffenburger, Fusing navigation and vision information with the Transferable Belief Model: Application to an intelligent speed limit assistant, Informat. Fusion (2013). http://dx.doi.org/10.1016/j.inffus.2013.05.013.
4. 2013 - Intelligent fire-detection model using statistical color models data fusion with Dezert-Smarandache method, Gurjit Walia, Ankit Gupta, Rajiv Kapoor, in International Journal of Image and Data Fusion, Vol. 4, No. 4, pp. 324--341, 2013.
5. 2013 - A DSmT based combination scheme for multi-class classification, N. Abbas, Y. Chibani, Z. Belhadi, M. Hedir, Proc. of Fusion 2013 Int. Conf., Istanbul, Turkey, July 9-12, 2013.
6. 2013 - Application of new absolute and relative conditioning rules in threat assessment, K. Krenc, F. Smarandache, Proc. of Fusion 2013 Int. Conf., Istanbul, Turkey, July 9-12, 2013.
7. 2013 - A evidential fusion network base context reasoning for smart media service, J. Yang, S. An, J.K. Choi, 15th Int. Conf on Advanced Communication Technology (ICACT), pp. 986-989, Jan 27-30, Phoenix Park, Pyeongchang, Korea, 2013.
8. 2013 - Multisensor data fusion: A review of the state-of-the-art, B. Khaleghi, A. Khamis, F.O. Karray, S.N. Razavi, Information Fusion, Volume 14, No. 1, Pages 28-44, January 2013.
9. 2013 - Multi-source information fusion for open innovation decision support system, by Li Li, Sun Lu, Wang Jiayang, Journal of Theoretical and Applied Information Technology, Vol. 50. No. 2, April 2013.
10. Li P., Huang X., Wang S., Dezert J., SLAM and path planning of mobile robot using DSmT, Journal of Software Engineering 7 (2), pp. 46-67, 2013.
11. 2013 - Using behavioral indicators to help detect potential violent acts - A review of the Science Base, P.K. Davis, W.L. Perry, R.A. Brown, D. Yeung, P. Roshan, P. Voorhies, RAND National Defense Research Institute report (RR215), ISBN 978-0-8330-8092-9, 2013.
12. 2013 - New method for diagnosing a motor with vibration fault, H. Zhan, J. Tan, Information Technology Journal, Vol. 12, Issue 8, pp. 1664-1667, 2013.
13. 2013 - A Dynamic Evidential Fusion Network for Decision Making in Multi- Robot System, H. Lee, International Journal of Information Processing and Management (IJIPM), Vol. 4, No. 3, pp. 156--170, 2013.
14. 2013 - Occupancy Grid Mapping Based on DSmT for Dynamic Environment Perception, Zhou Junjing, Duan Jianmin, Yang Guangzu, International Journal of Robotics and Automation (IJRA), Vol. 2, No. 4, pp. 129~139, December 2013.

15. 2013 - Application of DSmT-ICM with Adaptive decision rule to supervised classification in multisource remote sensing, A. Elhassouny, S. Idbraim, A. Bekkari, D. Mammass and D. Ducrot, Journal of computing, Vol. 5, No. 1, January 2013.
16. 2013 - Integrating textual analysis and evidential reasoning for decision making in Engineering design, F. Browne, N. Rooney, W. Liu, D. Bell, H. Wang, P.S. Taylor, Y. Jin, Knowledge-Based Systems, Vol. 52, pp. 165-175, 2013.
17. 2013 - Gearbox incipient fault fusion diagnosis based on DSmT and wavelet neural network, by Chen Fafa, Tang Baoping and Yao Jinbao, Journal of Vibration and Shock, Vol. 32, No. 9, pp. 40--45, 2013.
18. 2013 - Design on a Framework to Facilitate Decisions Using Information Fusion, by Tamer M. Abo Neama, Ismail A. Ismail, Tarek S. Sobh, M. Zaki, Journal of Al Azhar University Engineering Sector (JAUES), Vol. 8, No. 28, pp. 1237--1250, July 2013.

2012 Publications

1. 2012 - Pedestrian tracking using color, thermal and location cue measurements: a DSmT-based framework, M. Airouche, L. Bentabet, M. Zelmat, G. Gao, Machine Vision and Applications, Volume 23, Issue 5, pp 999-1010, September 2012.
2. 2012 - Two generalizations of aggregated uncertainty measure for evaluation of Dezert-Smarandache theory, Mahdi Khodabandesh, Alireza Mohammad-Shabri, International Journal of Information Technology & decision Making (IJITDM), Vol. 11, No. 1, pp. 119-142, 2012.
3. 2012 - Dezert J., Smarandache F., An introduction to DSmT for information fusion, Proc. of the Annual Symposium of the Institute of Solid Mechanics and Session of the Commission of Acoustics, Romanian Academy, edited by Luige Vladareanu and Marcel Migdalovici, pp. 13-24, The XXII-nd SISOM, 24-25 May 2011, Bucharest, Romania. Dezert J., Smarandache F., An introduction to DSmT for Information Fusion, World Scientific Publishing Company, New Mathematics and Natural Computation, Vol. 8, No. 3, pp. 343-359, 2012.
4. 2012 - A DSmT-Based Approach for Data Association in the Context of Multiple Target Tracking, M. Airouche, L. Bentabet, M. Zelmat, Intelligent Robotics and Applications, Lecture Notes in Computer Science Volume 7507, pp 686-695, 2012.
5. 2012 - Association of airborne target broken tracks based on DSmT, L. Gao, J. Jiang, J. Lan, Tranducer and Microsystem Technologies, April 2012.
6. 2012 - SVM-DSmT Combination for the simultaneous verification of Off-Line and On-Line handwritten signatures, N. Abbas, Y. Chibani , Int. J. Comp. Intel. Appl. Vol. 11, No. 3, 2012.
7. 2012 - SVM-DSmT Combination for Off-Line Signature Verification, Nassim Abbas and Youcef Chibani, International Conference on Computer, Information and Telecommunication Systems: CITS'12, Amman, Jordan, May 13 - 16, 2012.
8. 2012 - Handwritten Digit Recognition Based On a DSmT-SVM Parallel Combination, N. Abbas, Y. Chibani and H. Nemmour, 13th International Conference on Frontiers in Handwriting Recognition: ICFHR'12, Bari, Italy, September 18 - 20, 2012
9. 2012 - Change detection by new DSmT decision rule and ICM with constraints: Application to Argan land cover , A. Elhassouny, S. Idbraim, D. Mamass, D. Ducrot, International Conference on Multimedia Computing and Systems (ICMCS), pp. 107 – 112, 2012.
10. 2012 - Multisource Fusion/Classification Using ICM and DSmT with New Decision Rule, A. Elhassouny, S. Idbraim, A. Bekkari, D. Mamass, D. Ducrot, Proc. of 5th International Conference on Image and Signal Processing, ICISP 2012, Agadir, Morocco, June 28-30, 2012.

11. 2012 - DSmT Based RX Detector for Hyperspectral Imagery by H. Le, P. Zhang, W. Ruan, Symp. on Photonics and Optoelectronics (SOPO), 2012.
12. 2012 - Contribution of evidence-similarity to target classification, J. Li, J. Lan , 7th IEEE Conf. On Industrial Electronics and Applications (ICIEA), pp 1822-1826, 2012.
13. 2012 - An overview on the 2-tuple linguistic model for computing with words in decision making: Extensions, applications and challenges, L. Martinez, F. Herrera, Information Sciences, Volume 207, 10 November 2012, Pages 1-18.
14. 2012 - Diagnosis of aero-engine with early vibration fault symptom using DSmT, X. Zhai, J. Hu, S. Xie, J. Liu, Q. Li, Journal of Aerospace Power, February, 2012.
15. 2012 - DSmT based scheduling algorithm in OFDMA systems , Y. Ren, T. Lv, F. Long, J Zhang, Y. Lu, Telecommunications (ICT), 2012 19th International Conference on 2012 , Page(s): 1 – 6
16. 2012 - Updating Attribute Fusion Results with Additional Evidence Using DSmT, by K. Krenc, Proc. of the 15th International Conference on Information Fusion, Singapore, 9-12 July 2012.
17. 2012 - Comparative Study of Contradiction Measures in the Theory of Belief Functions, F. Smarandache, D. Han, A. Martin, Proceedings of the 15th International Conference on Information Fusion, Singapore, 9-12 July 2012.
18. 2012 - Neutrosophic Masses & Indeterminate Models. Applications to Information Fusion,F. Smarandache, Proc. of the 15th International Conference on Information Fusion, Singapore, 9-12 July 2012.
19. 2012 - New Mexico-Arizona Book Award at the category Science & Math, for the book "DSm Super Vector Space of Refined Labels", by W. B. Kandasamy, F. Smarandache, 16 November 2012, Albuquerque, NM, USA.
20. 2012 - Non-exclusive hypotheses in Dempster–Shafer Theory, Laurence Cholvy, International Journal of Approximate Reasoning, Volume 53, Issue 4, June 2012, Pages 493-501
21. 2012 - Plausibility in DSmT, M. Daniel, Conf . on Belief Functions: Theory and Applications Advances in Intelligent and Soft Computing Volume 164, 2012, pp 179-187.
22. 2012 - Threshold improvement method combining DSmT and DST, Y. Liu, S. Ling, Journal of Computer Applications, April 2012.
23. 2012 - A new DSmT combination rule in open frame of discernment and its application, C. Wen, X. Xu, H. Jiang, Z. Zhou, Science China Information Sciences, Volume 55, Issue 3, pp 551-557, March 2012.
24. 2012 - New conflict representation model in generalized power space , Y. He, L. Hu, X. Guan, D. Han Y. Deng, Journal of Systems Engineering and Electronics, Vol. 23 , No. 1, pp. 1-9, 2012.
25. 2012 - Neutrosophic masses & indeterminate models. Applications to information fusion , F. Smarandache, Int. Conf. On Advanced Mechatronic Systems (ICAMechS), pp. 674 – 679, 2012.
26. 2012 - Examination of combination rules for the purpose of information fusion in C2 systems: Application of Dezert-Smarandache Theory, Krenc K., Military Communications and Information Systems Conference (MCC) , 2012.
27. 2012 - Self-organization in an agent network: A mechanism and a potential application, D. Ye, M. Zhang, D. Sutanto, Decision Support Systems, Volume 53, Issue 3, June 2012, Pages 406-417.
28. 2012 - New metrics between bodies of evidences, P. Djiknavorian, D. Grenier, P. Valin, Journal of Emerging Technologies in Web Intelligence, Vol. 4, No. 3, pp. 264-272, Aug. 2012.
29. 2012 - Application of DSmT in facial expression recognition, C.M. Zao, J. Wei, Z.G. Xing, Z. Wei, Proc. of ASME 2012 International Mechanical Engineering Congress & Exposition, Houston, TX, USA, 9-15 Nov. 2012.

30. 2012 - One Fusion Approach of Fault Diagnosis Based on Rough Set Theory and Dezert-Smarandache Theory, Yanqin Su, Zhenyu Song, Tingxue Xu, in 2nd International Conference on Electronic & Mechanical Engineering and Information Technology (EMEIT-2012).
31. 2012 - Research of decentralized collaborative target tracking architecture in the sea battlefield for the complex sensor networks, Li Duan, Kun Feng, Bin Luo, Ya-Nan Li, Artificial Intelligence and computational intelligence, Lecture Notes in Computer Science, Vol. 7530, pp. 492-499, Springer, 2012.
32. 2012 - Cooperative Spectrum Sensing Using Discounted DSmT, by Zhang Jun-Nan; Zhang Shao-Wu, Signal Processing, Vol. 28, No. 2, pp. 166-171, Feb. 2012.
33. 2012 - Multi-scale fused edge detection algorithm based on conflict redistribution DSmT, by Qiao Kui-xian; Yin Shi-bai; Qu Sheng-jie, in Journal of Computer Applications, Vol. 32, No. 4, pp. 1050--1055, April 2012.
34. 2012 - DSmT based scheduling algorithm in OFDMA systems, by Yuan Ren; Tiejun Lv; Feichi Long; et al., in 19th International Conference on Telecommunications (ICT 2012), 23-25 April, 2012, Jounieh, Lebanon.

2011 Publications

1. 2011 - Measuring Conflict Functions in Generalized Power Space, L. Hu, X. Guan, Y. Deng, D. Han, Chinese Journal of Aeronautics, Volume 24, Issue 1, February 2011, Pages 65-73.
2. 2011 - Adaptive Information-Fusion Algorithm of DST and DSmT, Wang Qian, Wang Wei, Wu Minggui, Zou Changwen, Ordnance Industry Automation, 2011-02.
3. 2011 - Li X., Dezert J., Smarandache F., Huang X., Evidence supporting measure of similarity for reducing the complexity in information fusion, Information Sciences, Vol. 181, pp. 1818–1835, 2011.
4. 2001 - A composite Self-organization mechanism in an agent network, by D. Ye, M. Zhang, Q. Pai, Proc. of WISE 2011, Springer LNCS 6997, pp. 249-256.
5. 2011 - Change detection by fusion/contextual classification based on a hybrid DSmT model and ICM with constraints, Azeddine Elhassouny, Soufiane Idbraim, Aissam Bekkari, Driss Mammass and Danielle Ducrot, International Journal of Computer Applications 35(8):28-40, December 2011.
6. 2011 - Combination of Off-Line and On-Line Signature Verification Systems Based on SVM and DST, by Nassim Abbas and Youcef Chibani, 11th International Conference on Intelligent Systems Design and Applications (ISDA), Cordoba, Spain, November 22-24, 2011.
7. 2011 - Chinese translation of DSmT Book Vol. 1, by Prof. X. Li, June 2011.
8. 2011 - Interactive-Adaptive Combination Rule, Jin Hongbin, Lan Jiangqiao, Li Hongfei, China Communications, March, 2011.
9. 2011 - An Improved Focal Element Control Rule, H. Jin, J. Lan, Procedia Engineering, Volume 15, Pages 13-17, 2011.
10. 2011 - Application of evidential reasoning to improve the mapping of regenerating forest stands, B. Mora, R. Fournier, S. Foucher, International Journal of Applied Earth Observation and Geoinformation, Volume 13, Issue 3, June 2011, pp. 458-467.
11. 2011 - A novel approach for meeting the challenges of the integrated security systems, Moein Kasem, Nadim Chahin, WSEAS Transactions on Systems, Issue 9, Vol. 19, Sept. 2011.
12. 2011 - Improved DSmT and its application in radar type recognition,J. Yan, Y. Wang, X. Wang, Int. Conf. On Business Management and Electronic Information (BMEI), pp. 207-211, 2011.

13. 2011 - Research on Performance Evaluation for Selection of Supplier Based on DSmT, J. Li, Value Engineering Journal, No. 26, 2011.
14. 2011 - Approximate Importance Sampling Approach Based on DSmT in Structural Reliability Analysis, D. Chen, B. Jin, Proceedings of the GeoHunan International Conference II: Emerging Technologies for Design, Construction, Rehabilitation, and Inspection of Transportation Infrastructure, held in Hunan, China, June 9-11, 2011.
15. 2011 - Shedding New Light on Reason for Conflict between Evidences, L. Hu, J. Liu, X. Guan, P. Xin, W. Liu, T. Zhou, Int. Conf. On Network Computing and Information Security (NCIS), pp 405--409, 2011.
16. 2011 - A New Probabilistic Transformation in Generalized Power Space, L. Hu, Y. He, X. Guan, Y. Deng, D. Han,
17. Chinese Journal of Aeronautics, Volume 24, Issue 4, Pages 449-460, Aug. 2011.
18. 2011 - Research on Dynamic Targets Tracking Based on Color Cues Under Complicated Scene. Y. Wang, Y. Fang, X. Da, S. Chen, Procedia Engineering, Vol. 16, pp. 59-64, 2011.
19. 2011 - DSm Vector Spaces of Refined Labels, B.Kandasamy, F.Smarandache, Educ. Publ., Columbus, USA, 214 p., 2011.
20. 2011 - Recognition fusion based on DSmT with BP neural network , S. Shi, X. Wang, J. Lu, Int. Conf. On Computational Problem-Solving (ICCP), pp. 623 – 626, 2011.
21. 2011 - Multiple Moving Targets Tracking Research in Cluttered Scenes. Y. Fang, Y. Wang, W. Jin, Fang, International Workshop of Automobile, Power and Energy Engineering, Procedia Engineering, Vol. 16, p. 54-58, 2011.
22. 2011 - Research on Dynamic Targets Tracking Based on Color Cues Under Complicated Scene, Y. Wang X. Da, S. Chen, H. Wang, Procedia Engineering, International Workshop on Automobile, Power and Energy Engineering, Volume 16, Pages 59–64, 2011.
23. 2011 - Application of Referee Functions to the Vehicle-Born Improvised Explosive Device Problem, F. Dambreville, Proceedings of the 14th International Conference on Information Fusion, Chicago, Illinois, USA, July 5-8, 2011.
24. 2011 - Definition of bba and quality of fusion in C2 systems, K. Krenc, A. Kawalec, T. Pietkiewicz, Proceedings of the 14th International Conference on Information Fusion, Chicago, Illinois, USA, July 5-8, 2011.
25. 2011 - DSmT based scheduling algorithm in opportunistic beamforming systems, F. Long, T. Lv, H. Gao, P. Chang, 18th Int. Conf. On Telecommunications (ICT), pp. 534-538, 2011.
26. 2011 - Recognition fusion based on DSmT with BP neural network , S. Shi, X. Wang, J. Lu, Int. Conf. On Computational Problem-Solving (ICCP), pp. 623—626, 2011.
27. 2011 - Case retrieval in medical databases by fusing heterogeneous information, by G. Quellec, M. Lamard, G. Cazugue, C. Roux, B. Cochener, in IEEE Transactions on Medical Imaging 30, 1, 108-18, 2011.
28. 2011 - A novel hybrid method for mobile robot path planning in unknown dynamic environment based on hybrid DSm model grid map, Peng Li, Xinhan Huang, Min Wang, Journal of Experimental and Theoretical Artificial intelligence - Advances in knowledge discovery and data analysis for artificial intelligence, Vol. 23, No. 1, pp. 5-22, March 2011.

2010 Publications

1. 2010 - Smarandache F., Dezert J., An Algorithm for Quasi-Associative and Quasi-Markovian Rules of Combination in Information Fusion, Int. Journal of Applied Mathematics & Statistics, Vol. 22, No. S11 (Special Issue on Soft Computing), pp. 33-42, 2011. And BRAIN (Broad Research in Artificial Intelligence and Neuroscience) Journal, Vol. 1, Oct 2010 (Special Issue on Advances in Applied Sciences).
2. 2010 - Adaptive universal proportional redistribution rule under DSmT framework, P. Li, X. Huang, M. Wang, Computer Engineering and Applications, Vol.46 (6): 16-18, 2010.
3. 2010 - A Static Evidential Network for Context Reasoning in Home-Based Care , H. Lee, J. Choi, E. Ramez, IEEE Trans. On SMC, Part A, Vol. 40 , no. 6, pp 1232-1243, 2010.
4. 2010 - Fusion of communication interception information based on DSmT, H. Yang, K. Chen, J. Yang, Computer Engineering and Applications, Vol. 46 (10), pp.129-132, 2010.
5. 2010 - Fusion techniques for reliable information: A survey, Hyun lee, Byoungyong Lee, Kyungseo Park, Ramez Elmasri, International Journal of Digital Content Technology and its applications, Vol. 4, No. 2, April 2010.
6. 2010 - Analysis and research of Proportional Conflict Redistribution Rules in Identity Identification, K. Tian, H. Li, H. Jin, Proceedings of the 2nd IEEE International Conference on Future Computer and Communication (ICFCC 2010), Vol.1 pp 254-257, Wuhan, China, May 21-24th, 2010.
7. 2010 - A New Evidence Combination Method, H. Jin, J. Lan, H. Li, in Proceedings of the 2nd IEEE International Conference on Future Computer and Communication (ICFCC 2010), Vol.1 pp 382-385, Wuhan, China, May 21-24th, 2010.
8. 2010 - Union/Intersection vs. alternative/conjunction - defining posterior hypotheses in C2 systems, K. Krenc, A. Kawalec, Proceedings of Fusion 2010 International Conference, Edinburgh, Scotland, 26-29 July 2010.
9. 2010 - A new evidence combination method , H. Jin, J. Lan, H. Li, 2nd Int. Conf. On Future Computer and Communication (ICFCC), pp. V1-382 - V1-385, 2010.
10. 2010 - A method for condition evaluation based on DSmT , A. Liu, 2nd Int. Conf. On Information Management and Engineering (ICIME), pp. 263-266, 2010.
11. 2010 - The 6P's of Sales Resource Management , J. Wang, N. Lee, A. Timothy, Int. Conf. On Future Information Technology and Management Engineering (FITME), pp. 150-153, 2010.
12. 2010 - Implementation of Approximations of belief functions for fusion of ESM reports within the DSm framework, P. Djiknavorian, D. Grenier, P. Valin, Proceedings of Fusion 2010 International Conference, Edinburgh, Scotland, 26-29 July 2010.
13. 2010 - Unification of Fusion Rules (UFR), by Florentin Smarandache, in Multispace and Multistructure, Vol. 4, 285, 2010.
14. 2010 - Fusion of Masses Defined on Infinite Countable Frames of Discernment, F. Smarandache, A. Martin, in Multispace and Multistructure, Vol. 4, 299-303, 2010.
15. 2010 - A combination of ontology fusion and DSmT as the way to improve the performance of evidential inference, in C2 systems, K. Krenc, A. Kawalec, in Proceedings of Workshop on the theory of belief functions, April 1-2, 2010, Brest.
16. 2010 - Generic implementation of fusion rules based on Referee Function, in Proceedings of Workshop on the theory of belief functions, by F. Dambreville, April 1-2, 2010, Brest, France.
17. 2010 - Approximation in DSm theory for fusing ESM reports, P. Djiknavorian, D. Grenier, P. Valin, in Proceedings of Workshop on the theory of belief functions, April 1-2, 2010, Brest, France.

18. 2010 - A particle filtering and DSmT Based Approach for Conflict Resolving in case of Target Tracking with multiple cues, by Y. Sun and, L. Bentabet, Journal of Mathematical Imaging and Vision, Vol. 36, No 2, pp. 159-167, February 2010.
19. 2010 - An Intelligent Fusion Method of Sequential Images Based on Improved DSmT for Target Recognition , Z. Miao, Y. Cheng, Q. Pan, J. Hou, Z. Liu, Int. Conf. On Computational Aspects of Scial Networks (CASoN), pp. 369-373, 2010.
20. 2010 - Multi-sensor Fusion Based on DSmT and Particle Filtering, J. Xia, J. Yang, Q. Zhang, Computer Engineering, Vol.36 (20), pp. 179-181, 2010.
21. 2010 - A Realization of Auto Plotting Table Basing on MMHCI , J. Jin, W. Li, Int. Conf. On Machine Vision and Human-Machine Interface (MVHI), pp. 505-508, 2010.
22. 2010 - Analysis and improvement for Proportional Conflict Redistribution rules , H. Li, H. Jin, K. Tian, X. Fei, Computer Application and System Modeling (ICCASM), Page(s): V11-401 - V11-404, 2010.
23. 2010 - Fusion of qualitative beliefs based on linguistic labels , J. Gao, W. Lian, J. Yan, 2nd Int. Conf. On Signal Processing Systems (ICSPS), Page(s): V1-515 - V1-519, 2010.
24. 2010 - Modified combination rule based on DSmT, Qiao Qiansu, Electronic Measurement Technlogy Journal, February, 2010.
25. 2010 - Data Fusion of Cameras' Images for Localization of an Object; DSmT-based Approach, M. Khodabandeh, A. Mohammad-Shahri, The 1st International Symposium on Computing in Science & Engineering, 3-5 June 2010, Kusadasi, Turkey.
26. 2010 - Multispace and multistructures. Neutrosophic transdisciplinarity (100 collected papers of Sciences), Vol. IV, F. Smarandache (Editor), NESP, 2010, Chapter on Information fusion pp. 270-394.
27. 2010 - Fusion of possibilistic sources of evidences for pattern recognition. , by Dahabiah Anas, John Puentes, Basel Solaiman, in Integrated computer-aided engineering. April 2010, vol. 17, n° 2, pp. 117-130.
28. 2010 - Application of DSmT in integrated identification of friend-or-foe, byXin Yu-lin; Zou Jiang-wei; Xu Shi-you; et al., Systems Engineering and Electronics, Vol. 32, No. 11, pp. 2385-2388, Nov. 2010.
29. 2010 - A Fast Approximate Reasoning Method in Hierarchical DSmT, by Li Xin-de; Huang Xin-han; et al., Acta Electronica Sinica, Vo. 38, No. 11, pp. 2566-2572, Nov. 2010.
30. 2010 - Multi-sensor Fusion Based on DSmT and Particle Filtering, by Xia Jian-ming; Yang Jun-an; Zhang Qiong, in Computer Engineering, Vol. 36, No. 20, pp. 179-81, 184, Oct. 2010.
31. 2010 - High Complexity Optimization for DSmT Mixing Rule, by Chen Kai; Yang Jun-an; Chen Hao, in Computer Engineering, Vol. 36, No. 7, pp. 76-8, 81, April 2010.
32. 2010 - DSmT-based Mobile Robot Map Building and Sensor Management, by YANG Jin-yuan; HUANG Xin-han; LI Peng, in Computer Science Vol. 37, No. 4, pp. 227--230, 2010.
33. 2010 - Fusion of imprecise qualitative information, by Li X., Dai X., Dezert J., Smarandache F., Applied Intelligence vol. 33 (3), pp. 340--351, December 2010

2009 Publications

1. 2009 - Intelligent fusion algorithm of multi-source conflicting evidences, J. Yang, X. Huang, M. Wang, Computer Engineering and Applications, Vol. 45 (22), 2009.

2. 2009 - Design of an Intelligent Housing System Using Sensor Data Fusion Approaches, A.M. Khalkhali, B. Moshiri, H.R. Momeni, Chap. 11 in Sensor and Data Fusion, N. Milisavljevic (Editor), In-Tech Publisher, 2009. (here)
3. 2009 - Designing a home security system using sensor data fusion with DST and DSmT methods, A.M. Khalkhali, B. Moshiri, H.R. Momeni, IJE Trans. A: Basics, Vol. 22, No.1, February 2009 (here)
4. 2009 - A MES pre-processor design in multi-source information fusion , J. Yang, X. Huang, M. Wang, Yang, Jinyuan, Int. Conf on Cyber-Enabled Distributed Computing and Knowledge Discovery, CyberC '09 , pp. 331-336, 2009.
5. 2009 - A Hybrid Method for Dynamic Local Path Planning , P. Li, X. Huang, M. Wang, Int. Conf. On Networks Security, Wireless Communications and Trusted Computing, NSWCTC '09 , pp. 317-320, 2009.
6. 2009 - Qualitative reasoning by computing with words in Herrera-Martínez's linguistic model, X. Li, X. Dai, Journal of Electronics (China), Vol. 26, Issue 4, pp 564-570, July 2009.
7. 2009 - Study on the Threshold of Combining DST and DSmT, X. Zhou, Ship Electronic Eng. Journal, Dec. 2009.
8. 2009 - Evidence Reasoning Machine based on DSmT for mobile robot mapping in unknown dynamic environment, X. Huang, P. Li, M. Wang, IEEE Int. Conf. On Robotics and Biomimetics (ROBIO), pp. 753--758, 2009.
9. 2009 - The combination method for dependent evidence and its application for simultaneous faults diagnosis, H. Jiang, X. Xu, C. Wen, Int. Conf. On Wavelet Analysis and Pattern Recognition, 2009. ICWAPR 2009, pp. 496-501, 2009.
10. 2009 - Extension of Inagaki General Weighted Operators and A New Fusion Rule Class of Proportional Redistribution of Intersection Masses, by F. Smarandache, International Journal of Artificial Intelligence, Vol. 3, No. A09, 79-85, 2009.
11. 2009 - Application of DSM theory for SAR image change detection, S. Hachicha, F. Chaabane, 16th IEEE Int. Conf. On Image Processing (ICIP), pp. 3733 – 3736, 2009.
12. 2009 - Research on Moving Object Tracking under Occlusion Conditions , J. Sun, Y. Wang, Int. Conf on Management and Service Science, MASS '09, 2009.
13. 2009 - Using Logic to Understand Relations between DSmT and Dempster-Shafer Theory, L. Cholvy, ECSQARU 2009, pp. 264-274 (also in Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Lecture Notes in Computer Science Volume 5590, 2009, pp 264-274).
14. 2009 - User insisted redistribution of belief in hierarchical classification spaces, W. Van Norden, C.M. Jonker, in Proc. of the 2009 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Milan, Italy, September 15-18, 2009.
15. 2009 - An Efficient Fusion Algorithm on Conflicting Evidence, J. Yang, X. Huang, M. Wang, 8th IEEE/ACIS Int. Conf. On Computer and Information Science, 2009. ICIS 2009, pp. 650 – 654, 2009.
16. 2009 - A dynamic evidential network for multisensor context reasoning in home-based care, H. Lee, J. Choi, E. Ramez, IEEE Int. Conf on Systems, Man and Cybernetics, pp. 4994 – 4999, 2009.
17. 2009 - Sensor Data Fusion Using DSm Theory for Activity Recognition under Uncertainty in Home-Based Care, H. Lee, J. Choi, J. Sung, E. Ramez, Int. Conf. On Advanced Information Networking and Applications, AINA '09 pp. 517--524, 2009.
18. 2009 - DSm theory for fusing highly conflicting ESM reports, P. Valin, P. Djiknavorian, D.Grenier, Proceedings of Fusion 2009 International Conference, Seattle, USA, 2009.

19. 2009 - Modeling evidence fusion rules by means of referee functions, F. Dambreville, Proceedings of Fusion 2009 International Conference, Seattle, USA, 2009.
20. 2009 - Real world implementation of belief function theory to Detect Dislocation of Materials in Construction, S.N. Razavi, C.T. Haas, Ph. Vanheeghe, Emmanuel Duflos, Proceedings of Fusion 2009 International Conference, Seattle, USA, 2009.
21. 2009 - Dempster-Shafer Theory: combination of information using contextual knowledge, M.C. Florea, E. Bossé, Proceedings of Fusion 2009 International Conference, Seattle, USA, 2009.
22. 2009 - An application of DSmT in ontology-based fusion systems, K. Krenc, A. Kawalec, Proceedings of Fusion 2009 International Conference, Seattle, USA, 2009.
23. 2009 - Unification of Evidence Theoretic Fusion Algorithms: A Case Study in Level-2 and Level-3 Fingerprint Features, M. Vatsa, R. Singh, and A. Noore, IEEE Transactions on Systems, Man, and Cybernetics - A , Vol 29, No. 1, 2009.
24. 2009 - Applying the PCR6 Rule of Combination in Real Time Classification Systems, K.A. Scholte, W. Van Norden, Proceedings of Fusion 2009 International Conference, Seattle, USA, 2009.
25. 2009 - Implication of Culture: User Roles in Information Fusion for Enhanced Situational Understanding, E. Blasch, P. Valin, E. Bossé, M. Nilsson, J. Van Laere, E. Shahbazian, Proceedings of Fusion 2009 International Conference, Seattle, USA, 2009.
26. 2009 - An improved partial Haar dual adaptive filter for rapid identification of a sparse echo channel, P. Kechichian, B. Champagne, Signal Processing, Volume 89, Issue 5, pp. 710-723, May, 2009.
27. 2009 - A new evidential trust model for open communities, J . Wang, H. Sun, Computer Standards & Interfaces, Volume 31, Issue 5, September 2009, Pages 994-1001, Sept. 2009.
28. 2009 - Robust combination rules for evidence theory, M.C. Florea, A.-L. Jousselme, É. Bossé, D. Grenier, Information Fusion, Vol. 10, No. 2, pp. 183-197, April 2009.
29. 2009 - Quality-augmented fusion of level-2 and level-3 fingerprint information using DSm theory, M.Vatsa, R. Singh, A. Noore, M. Houck, International Journal of Approximate Reasoning, Vol. 50, No. 1, pp. 51-61, January 2009.
30. 2009 - A new approach of simultaneous faults diagnosis based on random sets and DSmT, Z. Li, X. Xu, Journal of Electronics (China), Vol. 26, No. 1, pp 24-30, January 2009.
31. 2009 - An approximate reasoning method in Dezert-Smarandache Theory, X. Li, X. Wu, J. Sun Z. Meng, Xinde Li, Journal of Electronics (China), Vol. 26, No. 6, pp 738-745, November 2009.
32. 2009 - Multimodal Medical Case Retrieval using Dezert-Smarandache Theory with a priori knowledge, G. Quellec, M. Lamard, G. Cazuguel, B. Cochener, C. Roux, 4th European Conference of the International Federation for Medical and Biological Engineering, IFMBE Proceedings Vol. 22, pp 716-719, 2009.
33. 2009 - Expert referral model based on DSmT, J. Wang, T. Zhang, H. Sun, Computer Engineering, Vol. 35, No. 8, 2009.
34. 2009 - Conflict-redistribution DSmT and new methods dealing with conflict among evidences, S. Qu, Y. Cheng, Q. Pan, Y. Liang, S. Zhang, Control and Decision Journal, December 2009.
35. 2009 - Study on environment perception of mobile robots using DSmT-based fusion machine, X. Li, X. Huang, X. Dai, Z. Meng, Journal of Huazhong University of Science and Technology (Nature Science Edition), December 2009.
36. 2009 - Model for Dependent Evidences in DSmT Framework, by WANG Jin; SUN Huai-Jiang, in Computer Science Vol. 36, No. 8, pp. 260-263,295, 2009.
37. 2009 - A New Combination Rule based on DSmT, by HU Li-fang; GUAN Xin; HE You, in Fire Control and Command, Vol. 34, No. 7, pp. 9-11, 2009.

2008 Publications

1. 2008 - A new tool applied to robot perception by selecting evidence sources , H. Wen, X. Huang, X. Li, IEEE Int. Conf. On Automation and Logistics, ICAL 2008, pp. 2364--2369, 2008.
2. 2008 - Preceding car tracking using belief functions and a particle filter, J. Klein, C. Lecomte, P. Miché, IEEE International Conference on Pattern Recognition (ICPR'08), Tampa (USA), pp. 864–871, December 2008.
3. 2008 - Efficient combination rule of Dezert-Smarandache theory, L. Hu, X. Guan, Y. He, Journal of Systems Engineering and Electronics, Vol. 19, No. 6, pp. 1139-1144, December 2008.
4. 2008 - Extension of Inagaki General Weighted Operators and A New Fusion Rule Class of Proportional Redistribution of Intersection Masses, F. Smarandache, presented as poster at SWIFT 2008 - Skovde Workshop on Information Fusion Topics, Sweden.
5. 2008 - Efficient combination rule of Dezert-Smarandache theory , L. Hu, X. Guan, Y. Hou, Journal of Systems Engineering and Electronics, Vol. 19 , No. 6 , pp.1139--1144, 2008.
6. 2008 - Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition, R. Singh, M.Vatsa, A. Noore, Pattern Recognition, Vol. 41, No. 3, pp. 880-893, March 2008.
7. 2008 - A New Approach of Simultaneous Faults Diagnosis Based on DSmT, Z. Li, X. Xu, C. Wen, Journal of Hangzhou Dianzi University, June 2008.
8. 2008 - Multiple Mobile robots Map Building Based on DSmT , P. Li, X. Huang, M. Wang, X. Zeng, IEEE Int. Conf on Robotics, Automation and Mechatronics, pp. 509--514, Sept. 2008.
9. 2008 - Contribution of DSm approach to belief function theory, M. Daniel, Proc. of IPMU'08.
10. 2008 - Multimodal Medical Case retrieval using Bayesian networks and the Dezert-Smarandache Theory, by G. Quellec, M. Lamard, L. Bekri, G. Cazuguel, C. Roux, B. Cochener, in the Fifth IEEE International Symposium on Biomedical Imaging (ISBI '08), Paris, France, June 14-17, 2008,
11. 2008 - A Sequential Monte-Carlo and DSmT Based Approach for Conflict Handling in case of Multiple target Tracking, Y. Sun, L. Bentabet, Technical Report, 2008-003, Bishop's University, Canada.
12. 2008 - Combining System and User Belief on Classification Using the DSmT Combination Rule, W. Van Norden, Fusion 2008 International Conference, June 30-July 3, 2008, Cologne, Germany.
13. 2008 - An Evaluation of the Attribute Information for the Purpose of DSmT Fusion in C&C Systems, K. Krenc, A. Kawalec, Fusion 2008 International Conference, June 30-July 3, 2008, Cologne, Germany.
14. 2008 - Multiple camera fusion based on DSmT for tracking objects on ground plane , E.O. Garcia, L. Altamirano, Fusion 2008 International Conference, June 30-July 3, 2008, Cologne, Germany.
15. 2008 - Discrete Labels and Rich Foci in Theory of Evidence, C. Osswald, A. Martin, Fusion 2008 International Conference, June 30-July 3, 2008, Cologne, Germany.
16. 2008 - Combining system and user belief on classification using the DSmT combination rule , W. Van Norden, F. Bolderhei, C. Jonker, Information Fusion, 2008 11th International Conference on 2008.
17. 2008 - Multi-sensor Target Identification Based on DSmT, L. Hu, X. Guan, Y. He, Journal of Projectiles,Rockets,Missiles and Guidance, February 2008.
18. 2008 - System and method for combining diagnostic evidences for turbine engine fault detection, by V. Guralnik, D. Mylaraswamy, H.C. Voges, U. S. Patent 7337086, Honeywell Int. Inc., Feb, 2008.

19. 2008 - A fuzzy qualitative framework for connecting robot qualitative and quantitative representations, H. Liu, IEEE Trans. Fuzzy Systems, Vol. 16, No. 6, pp. 1522–1530, 2008.
20. 2008 - Approches de classification évidentielles et paradoxale d'images satellitaires multispectrales pour l'amélioration de la carte d'occupation du sol : Application au milieu urbain et periurbain de la région d'Alger, R. Khedam, A. Bouakache, G. Mercier, A. Belhadj-Aissa, Revue française de photogrammétrie et de télédétection, No190, pp. 28-39, 2008.

2007 Publications

1. 2007 - A sequential Monte Carlo and DSmT based approach for conflict handling in case of multiple targets tracking, Y. Sun, L. Bentadet, Lecture Notes in Computer Science 4633, pp. 526–537, 2007.
2. 2007 - Inverse Problem in DSmT and Its Applications in Trust Management, J. Wang, H. Sun, in The First International Symposium on Data, Privacy, and E-Commerce, ISDPE 2007, 1-3 Nov. 2007.
3. 2007 - Empirical analysis of generalized uncertainty measures with Dempster Shafer fusion, by Pong, Peter, Subhash Challa, in Int. Conf. on Information Fusion 2007, July 2007, Québec, Canada.
4. 2007 - Decision Level Multiple Cameras Fusion Using Dezert-Smarandache Theory, E. Garcia, L. Altamirano, in Lectures Notes in Computer Sciences, Springer Verlag 2007.
5. 2007 - Analyzing the combination of conflicting belief functions, Philippe Smets, Information Fusion, Vol. 8, No. 4, pp. 387-412, October 2007.
6. 2007 - Toward a combination rule to deal with partial conflict and specificity in belief functions theory, by A. Martin, and C. Osswald, in 10th International Conference on Information Fusion, 2007, Québec City, July 2007.
7. 2007 - Classical Belief Conditioning and its Generalization to DSm Theory, by Milan Daniel, Proceedings of the Sixth International Conference on Information and Management Sciences, San Luis Obispo, California Polytechnic State University, USA, edited by Lee, T., Liu, Y., Zhao, X., pp 596-603, 2007.
8. 2007 - Designing a home security system using sensor data fusion with DST and DSmT methods , B. Moshiri, A. Khalkhali, A. Moussavi, H. Momeni, Proc. Of Fusion 2007. (here)
9. 2007 - Reasoning under uncertainty: from Bayesian to Valuation Based Systems. Application to target classification and threat evaluation, book by A. Benavoli, L. Chisci, B. Ristic, A. Farina, A. Graziano; Università degli Studi di Firenze, Selex / Sistemi Integrati, Italy, September 2007.
10. 2007 - On the combination and normalization of interval-valued belief structures, Y. Wang, J. Yang, D. Xu, K. Chin, Information Sciences, Vol. 177, No. 5, Pages 1230-1247, March 2007.
11. 2007 - La Teoria di Dezert-Smarandache [in Italian], book chapter in "Il vero e il plausibile", by Emiliano Ippoliti, Morrisville, USA, 2007.
12. 2007 - Unification of Evidence Theoretic Fusion Algorithms: A Case Study in Level-2 and Level-3 Fingerprint Features, by M. Vatsa, R. Singh, and A. Noore, Proceedings of IEEE Conference on Biometrics: Theory, Applications and Systems, 2007.
13. 2007 - Robot map building based on fuzzy-extended DSmT, X. Li, X. Huang, Z. Wu, G. Peng, M. Wang, Y. Xiong, Proceedings of SPIE, the International Society for Optical Engineering, International Symposium on Multispectral Image Processing and Pattern Recognition No5, vol. 6787, pp. 67870Y.1-67870Y.9, Wuhan, China, 2007.

14. 2007 - Analysis of Information Fusion Combining Rules under the DSm Theory using ESM Inputs, P. Djiknavorian, D. Grenier, P. Valin, Fusion 2007 Int. Conf. on Information Fusion, Québec City, July 2007.
15. 2007 - Mobile Robot's Map Reconstruction Based on DSmT and Fast-Hough Self-Localization X. Huang, X. Li, Z. Wu, M. Wang, Int. Conf. On Information Acquisition, ICIA '07, pp. 590—595, 2007.
16. 2007 - The DSm approach as a special case of the Dempster-Shafer theory, Milan Daniel, proc. of ECSQARU'07, 2007.
17. 2007 - A Dependent Evidence Model of DSmT and the Approximate Solution of the Inverse Problem, by WANG Jin; SUN HuaiJiang, in Computer Science, Vol. 34, No. 9, pp. 200-202,244 , 2007.

2006 Publications

1. 2006 - Solving Information Fusion Problems on Unreliable Evidential Sources with Generalized DSmT, L. Cheng, L. Kong, X. Li, Journal of Applied Science, vol. 6, Issue 7, p.1581-1585, 2006.
2. 2006 - Target Recognition a Method of Sequential Images based on the Weight-DSmT, J. Hou, Z. Miao, Q. Pan, Fire Control and Command Control, July 2006.
3. 2006 - Generalization of the classic combination rules to DSm hyper-power sets, Milan Daniel, Info & Security, An International Journal, Vol. 20, pp. 50--64, 2006.
4. 2006 - DSmT Coupling with PCR5 for Mobile Robot's Map Reconstruction, X. Li, X. Huang, M. Wang, J. Xu, H. Zhang, Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation, pp. 887--892, 2006.
5. 2006 - A Adaptive Integration Algorithms with DST and DSmT, J. Hou, Z. Miao, Q. Pan, Microelectronics & Computer, October 2006.
6. 2006 - Intelligent target recognition method of sequential images based on DSmT, J. Hou, Z. Miao, Q. Pan, Journal of Computer Applications, January 2006.
7. 2006 - Prise en compte de l'incertitude dans une démarche de modélisation prédictive : Le cas de la LGV PACA, R.M. Basse, RTP, MoDyS, 5 p.
8. 2006 - Ordered DSmT and Its Application to the Definition of Continuous DSm Models, F. Dambreville, Information & Security, An International Journal, pp. 85-103, Vol. 20, 2006.
9. 2006 - Understanding the large family of Dempster-Shafer theory's fusion operators-a decision-based measure, C. Osswald, A. Martin, 9th international conference on information fusion, Florence, Italy, 2006.
10. 2006 - Robot Map Building from Sonar Sensors and DSmT, X. Li, X. Huang, M. Wang, Information & Security, An International Journal, pp. 104-121, Vol. 20, 2006.
11. 2006 - Improvement of Land Cover Map from Satellite Imagery using DST and DSmT, R. Khedam, A. Bouakache, G. Mercier, A. Belhadj-Aissa, Information and Communication Technologies, ICTTA '06, 2006.
12. 2006 - Human Expert Fusion for Image Classification, A. Martin, C. Osswald, Information & Security, An International Journal, pp. 122-143, Vol. 20, 2006.
13. 2006 - A Comparison of the Effect of Sonar Grid Map Building Based on DSmT and DST, X. Li, X. Huang, M. Wang, G. Peng, The Sixth World Congress on Intelligent Control and Automation,WCICA 2006, pp. 4073--4077, 2006.
14. 2006 - Sonar Grid Map Building of Mobile Robots Based on DSmT, X. Li, X. Huang, M. Wang, Information Technology Journal, Vol. 5, no 2, pp. 267-272, 2006.

15. 2006 - Robot map building from sonar and laser information using DSmT with discounting theory, Li, X., Huang X., Wang M., in Int. J. Inf. Technol., Vol. 3, No.2, 2006.
16. 2006 - An adaptation of DST and DSmT to design supervised multispectral classifiers, R. Khedam, A. Bouakache, G. Mercier, A. Belhadj-Aissa, Proc. of the 2nd Int. Symposium on Recent Advances in Quantitative Remote Sensing: RAQRS'2, Torrent (Valencia), Spain, 25-29th Sept. 2006.
17. 2006 - An approach for partner selection in supply chains based on weighted-DSmT, by Huang M. et al., in Conference: 8th West Lake International Conference on Small and Medium Business, pp. 1404-1409, Hangzhou, China, Oct. 15--17, 2006.

2005 Publications

1. 2005 - Rapport Fusion de Données, Rapport Technique, Marie-Lise Gagnon, Départ. de Génie Electrique et de Génie Informatique, Université Laval, Canada, November 2005.
2. 2005 - Example of Continuous Basic Belief Assignment in the DSmT Paradigm, F. Dambreville, Proceedings of Fusion 2005 Int. Conf. on Information Fusion, Philadelphia, PA, USA, July 25-29, 2005.

2004 Publications

1. 2004 - Vagueness, a multifacet concept - a case study on Ambrosia artemisiifolia predictive cartography, P. Maupin, A.-L. Jousselme, A.-L., in IEEE International Geoscience and Remote Sensing Symposium, IGARSS '04. Proceedings, Volume: 1, 20-24 Sept. 2004.
2. 2004 - Unification of Fusion Theories (UFT), Florentin Smarandache, International Journal of Applied Mathematics & Statistics, Roorkee, India, Vol. 2, 1-14, 2004.

Seminars on DSmT

2014

1. Advances and Applications of DSmT in Information Fusion, by J. Dezert, F. Smarandache; presented by F. Smarandache, Osaka University, Inuiguchi Laboratory, Department of Engineering Science, Japan, 10 January 2014.
2. alpha-Discounting Method for Multicriteria Decision Making, by F. Smarandache, Osaka University, Inuiguchi Laboratory, Department of Engineering Science, Japan, 10 January 2014.
3. Recent advances in information fusion using belief functions, by J. Dezert, School of Automation, Northwestern Polytechnical University, Xi'an, China, November 24th, 2014.
4. Recent advances in information fusion using belief functions, by J. Dezert, Southeast University, Nanjing, China, November 25th, 2014.

2013

1. Advances and Applications of DSmT in Information Fusion, by J. Dezert, F. Smarandache; presented by F. Smarandache, Okayama University of Science, Kroumov's Laboratory, Department of Intelligent Engineering, Okayama, Japan, 17 December 2013.
2. Advances of Dezert-Smarandache Theory in Information Fusion and its Applications to Medicine and Cybernetics, by J. Dezert and F. Smarandache; presented by F. Smarandache, Ohio State University, Wexner Medical Center, Department of Radiology, Columbus, Ohio, USA, 22 February 2013.
3. A new BBA approximation technique, and applications of belief functions for edge and change detection in images, by J. Dezert, Seminar given at Dept of Automation, School of Electronic and Information Engineering, Jiaotong University, Xi'an, July 23th, 2013.
4. Introduction to DST and DSmT - Advances on Information Fusion using belief functions, Invited lecture at Northwestern Polytechnical University, Xi'an, July 29th, 2013.
5. Advances in Information fusion and decision-making using belief functions, by J. Dezert, International Workshop on Information Fusion (IWIF2013), Jiaotong University, Xi'an, July 30th, 2013.
6. Advances in information fusion based on belief functions, by J. Dezert, IEEE (Section of Québec) Invited Lecture, Département de Génie Electrique et de Génie Informatique, Université Laval, Québec, Canada, November 14th, 2013.

2012

1. A fundamental contradiction in Dempster-Shafer Theory, by J. Dezert, (with A. Tchamova as co-author), Panel Discussion on Uncertainty, Fusion 2012, July, Singapour. (invited)

2011

1. DSmT in Information Fusion, by J. Dezert, F. Smarandache, presented by F. Smarandache, Northwest Polytechnic University, Institute of Control and Information, Xi'an, China, December 27, 2011.
2. Advances and Applications of DSmT for Information Fusion, by J. Dezert & F. Smarandache, presented by Florentin Smarandache, Institute of Atomic Physics, Măgurele, Romania, 26 May 2011; seminar organized by Dr. Marian Apostol.
3. Advances and Applications of DSmT for Information Fusion, by J. Dezert & F. Smarandache, presented by Florentin Smarandache, Romanian Academy, Institute of Solid Mechanics and Commission of Acoustics, Bucharest, 25 May 2011.
4. On the inconsistency of Dempster's rule of combination with the probability calculus, by J. Dezert, Journée Fusion, Télécom Bretagne, Brest, France, Nov. 17th, 2011. (invited)
5. DSmT in Information Fusion, by J. Dezert, F. Smarandache, presented by F. Smarandache, Northwest Polytechnic University, Institute of Control and Information, Xi'an, China, December 27, 2011.

2010

1. Avancées en fusion de données: Approche par DSmT, by Jean Dezert, Télécom Bretagne, Brest, France, October 8th, 2010.
2. Threat Assessment of a Possible Vehicle-Borne Improvised Explosive Device using DSmT, by J. Dezert, F. Smarandache, Forum of Fusion 2010 International Conference, Edinburgh, Scotland, 26-29 July, 2010.
3. Advances and Applications of DSmT for Information Fusion, by J. Dezert & F. Smarandache, presented by F. Smarandache, ENSIETA, Brest, France, 17 June 2010.
4. An Introduction to Information Fusion Level 1 and to Neutrosophic Logic/Set with Applications, ENSIETA (National Superior School of Engineers and the Study of Armament), Brest, France, 2 July 2010.

2009

1. Advances and Applications of DSmT for Information Fusion (third version), by F. Smarandache (co-author J. Dezert), Invited speaker at and sponsored by Air Force Institute of Technology Wright-Patterson AFB in Dayton, Ohio, USA, May 1st, 2009.
2. Fusion of Sensors' Information in the Machine Building Industry (DSmT), by F. Smarandache (co-author J. Dezert), Universitatea din Craiova, Facultatea de Mecanică, Romania, 18 May 2009.
3. Advances in Quantitative and Qualitative Information Fusion, by J. Dezert (co-author F. Smarandache), Lecture given at Institute of Intelligent Control School of Automation Southeast University, Nanjing, China, May 19th, 2009.
4. Some Applications of Quantitative and Qualitative Information Fusion, by Jean Dezert (co-author F. Smarandache), Lecture given at Institute of Intelligent Control School of Automation Southeast University, Nanjing, China, May 20th, 2009.
5. Introduction to DSmT for Information Fusion, by Jean Dezert (co-author F. Smarandache), Lecture given Huazhong University of Sciences and Technology, Wuhan, China, May 23rd, 2009.
6. Quantitative and Qualitative Information Fusion, by Jean Dezert (co-author F. Smarandache), International Workshop on Information Fusion, Lecture given at Beihang University, Beijing, China, May 28th, 2009.
7. Quantitative and Qualitative Information Fusion, by Jean Dezert (co-author F. Smarandache), International Workshop on Information Fusion, Lecture given at Xi'an Jiaotong University, Xi'an, China, May 31st, 2009.
8. Quantitative and Qualitative Information Fusion, by Jean Dezert (co-author F. Smarandache), International Workshop on Information Fusion, Lecture given at Hangzhou Dianzi University, Hangzhou, China, June 4th, 2009.
9. Advances and Applications of DSmT for Information Fusion (third version), by F. Smarandache (co-author J. Dezert), Air Force Research Laboratory (AFRL), Rome, NY, USA, June 29, 2009.
10. Advances and Applications of DSmT for Information Fusion (third version), by F. Smarandache (co-author J. Dezert), Griffiss Institute, Rome, NY, USA, July 1st, 2009.
11. Advances and Applications of DSmT for Information Fusion (fourth version), by J. Dezert & F. Smarandache, Fusion 2009 International Conference, Seattle, USA, July 6th, 2009.
12. An Introduction to Fusion Level 1 and to Neutrosophic Logic/Set with Applications, presented by F. Smarandache at Air Force Research Laboratory, in Rome, NY, USA, July 29, 2009.

2008

1. Extension of Inagaki General Weighted Operators and A New Fusion Rule Class of Proportional Redistribution of Intersection Masses, by F. Smarandache, presented as poster at SWIFT 2008 - Skovde Workshop on Information Fusion Topics, Sweden.
2. Advances and Applications of DSmT for Information Fusion, Tutorial by J. Dezert & F. Smarandache, Fusion 2008 International Conference, June 30-July 3, 2008, Cologne, Germany.
3. Applications de la DSmT en pistage et robotique, by J. Dezert, Seminar at Institute Henri Poincaré, 11 Rue Pierre et Marie Curie, 75005 Paris, France, 22 January, 3-4 p.m., 2008; http://lastre.asso.fr/aubin/Sem-Viab-cont.html

2007

1. Advances and Applications of DSmT for Information Fusion, by J. Dezert, F. Smarandache, Tutorial at Fusion 2007 Int. Conference on Information Fusion, Québec City, July 2007.
2. Introduction à la Fusion et au Conditionnement de Croyances Quantitatives ou Qualitatives dans le Cadre de la DSmT, by J. Dezert, Seminar on "Recents developpements en fusion de données", ONERA, Chatillon, France, June 14th, 2007.
3. Advances and Applications of DSmT for Information Fusion, Pre-Conference Workshop, by J. Dezert, F. Smarandache, Sensor Fusion Europe (Marcus Evans), 2nd Annual Assuring Accessibility of Secure Information for the Warfighter, Brussels, Belgium, January 29, 2007.

2006

1. Advances and Applications of DSmT for Plausible and Paradoxical Reasoning for Information Fusion, by J. Dezert, International Workshop organized by the Bulgarian IST Centre of Competence in 21st Century, December 14, 2006, Bulg. Acad. of Sciences, Sofia, Bulgaria.
2. Fusion of Qualitative and Quantitative Information using DSmT, Tutorial MO2 by J. Dezert, F. Smarandache, Fusion 2006 International Conference, Florence, Italy, July 10-13, 2006.
3. Fusion d'informations incertaines et conflictuelles par la DSmT, by Jean Dezert, Round panel Discussion on Prevision Methods, 38ièmes Journées de Statistique du 29 Mai 2006, EDF Recherche et Developpement (ICAME/SOAD), Clamart (92), France; http://www.jds2006.fr/index.php.
4. DSmT Approach for Quantitative and Qualitative Information Fusion, Invited Lecture given by F. Smarandache (co-author J. Dezert), University Kristen Satya Wacana, Salatiga, May 26, 2006, Indonesia.
5. DSmT Approach for Quantitative and Qualitative Information Fusion, Invited Lecture given by F. Smarandache (co-author J. Dezert), University Sekolah Tinggi Informatika & Komputer Indonesia, Malang, May 20, 2006, Indonesia.
6. DSmT Approach for Quantitative and Qualitative Information Fusion, Invited Lecture given by Jean Dezert in Information Days on Advanced Computing, supported by EU BIS21++ project and organised by Institute for Parallel Processing, Bulgarian Academy of Sciences, March 21-23, 2006, Hotel "Kamena", Velingrad, Bulgaria.

2005

1. DSmT: Une Nouvelle Approche pour la Gestion d'Informations Conflictuelles, Jean Dezert, ENSIETA, Brest, France, December 15, 2005, http://www.ensieta.fr/e3i2/Seminaire.html.
2. Dezert-Smarandache Theory (DSmT) of Plausible and Paradoxical Reasoning, J. Dezert, F. Smarandache, Sensor Fusion Europe, Post-Conference Workshop, Marcus Evans Co., November 4th, 2005, Barcelona, Spain (invited).

2004

1. To Be and Not To Be - An Introduction to Neutrosophic: A Novel Decision Paradigm, F. Smarandache, S. Bhattacharya, Jadavpur University, Kolkata, India, 23 December 2004.
2. Uvod Do Dezert-Smarandachovy Teorie (domnenkove funkce, zobecnene domnenkove funkce a jejich kombinovani) [About Dezert-Smarandache Theory (belief functions)], M. Daniel, Czech Society for Cybernetics and Informatics, Praha, Czech Republic, 7 December, 2004.
3. An In-Depth Look at Information Fusion Rules & the Unification of Fusion Theories, by Florentin Smarandache, Presented at NASA Langley Research Center (Hampton, Virginia), on November 5th, 2004.
4. Recent Applications of Dezert-Smarandache Theory for Information Fusion, J. Dezert, A. Tchamova, T. Semerdjiev, P. Konstantinova, S. Corgne, L. Hubert-Moy, M. Grégoire, NASA Langley Research Center, Hampton, VA, USA, 5 November 2004.
5. An Overview of DSmT for Information Fusion, J. Dezert, F. Smarandache, NASA Langley Research Center, Hampton, VA, USA, 5 November 2004; http://www.nianet.org/ecslectureseries/smarandache_110504.php, http://www.nianet.org/ecslectureseries/dezert_110504.php.
6. Nouvelles Avancées en Fusion d'Information, J. Dezert, MAnifestation des JEunes Chercheurs STIC, Université du Littoral, Côte d'Opale, Calais, France, 14 October 2004, http://lil.univ-littoral.fr/majecstic/.
7. Panel Discussion on DSmT Fusion 2004 Int. Conf. on Information Fusion, June 28-July 1, 2004, Stockholm, Sweden, http://www.fusion2004.org.
8. Introduction à la DSmT pour la fusion de données incertaines et conflictuelles by Jean Dezert, Seminar given at the Direction Générale de l'Armement (DGA/CTA/DT/GIP/PRO), 16 bis, Av. Prieur de la Côte d'Or, 94114 Arcueil Cedex, France, January 20th, 2004.

2003

1. A new framework for data fusion based on DSmT by Jean Dezert, Seminar given at Dept. of Elec. Eng., Univ. of Melbourne, Melbourne, Australia, July 24th, 2003.
2. Théorie du raisonnement plausible et paradoxal pour la fusion d'informations by Jean Dezert, Seminar given at Laboratoire PSI - FRE CNRS 2645, Université de Rouen, UFR des Sciences, Place Emile Blondel, 76821 Mont Saint Aignan Cedex, France, April 10th, 2003.
3. Théorie du raisonnement plausible et paradoxal pour la fusion d'informations by Jean Dezert, Seminar given at Laboratoire COSTEL, LETG UMR CNRS 6554, Université Rennes 2 Haute-Bretagne, Rennes, France, January 17th, 2003.

2002

1. Théorie du raisonnement plausible et paradoxal pour la fusion d'informations, by Jean Dezert, Seminar given at Institute Henri Poincaré,11 rue Pierre et Marie Curie 75005 - Paris, France, November 12th, 2002.

M.Sc. and Ph.D Theses related with DSmT

1. 2004 – France - Ph.D. Thesis, Modélisation prédictive de l'occupation des sols en contexte agricole intensif : application à la couverture hivernale des sols en Bretagne, Ph. D. Dissertation, by Samuel Corgne, Advisor: Laurence Hubert-Moy, Université de Rennes 2, France, 10 December 2004. (in French)
2. 2004 – Australia - Ph.D. Thesis, Utility, Rationality and Beyond - From Behavioral Finance to Informational Finance, Ph. D. Dissertation, by Sukanto Bhattacharya, Bond University, Queensland, Australia, 2004.
3. 2006 – Iran - M.Sc. Thesis, Designing a Home Security System Using Sensor Data Fusion, Arezou Moussavi, Islamic Azad University, Southern Unit, Tehran, Iran, September 2006.
4. 2007 – Canada - Ph.D. Thesis, Combinaison d'informations hétérogènes dans le cadre unificateur des ensembles aléatoires : Approximations et robustesse, by Mihai Cristian Florea, Université Laval, Québec City, Canada, 13 July 2007.
5. 2006 – China - Ph.D. Thesis, The Combination Rules、Performance Indexes and Applications of Evidence Reasoning, J. Hou, Northwestern Polytechnical University.
6. 2007 – China – Ph.D Thesis, Research on fusion method of imperfect information from multi-source and its application, by X.D. Li, Huazhong University of Science and Technology, China, 14 June 2007.
7. 2007 – China – Ph.D. Thesis, Information Superiority Oriented Study on Key Techniques of C-4ISR System, Y. Wang, Northwestern Polytechnical University, May 2007.
8. 2008 – Canada - M.Sc. Thesis, Fusion d'informations dans un cadre de raisonnement de Dezert-Smarandache appliquée sur des rapports de capteurs ESM sous le STANAG 1241, Pascal Djiknavorian, Laboratoire de Radiocommunication et de Traitement du Signal, Laval University, Canada, September 19, 2008.
9. 2008 – Italy - Ph.D. Thesis, Modelling and efficient fusion of uncertain information: beyond the classical probability approach, Alessio Benavoli, Ph.D. in Computer and Control Engineering, University of Firenze, Firenze, Italy, April, 2008.
10. 2008 – USA - Ph.D. Thesis, Quality Induced Secure Multiclassifier Fingerprint Verification using Extended Feature Set, by Mayank Vatsa, Computer Sciences Dept., West Virginia University, Morgantown, West Virginia, USA, Nov. 2008.
11. 2008 – France - Ph.D. Thesis, Indexation et fusion multimodale pour la recherche d'information par le contenu. Application aux bases de données d'images médicales, by G. Quellec, Ecole Nationale Supérieure des Telecommunications de Bretagne, Brest, France, September 19th, 2008.
12. 2009 – China - M.Sc. Thesis, Information Fusion Method Based on Dezert-Smarandache Theory for Fault Diagnosis, by H. Jiang, Hangzhou University of Electronic Science and Technology.
13. 2009 – Algeria - MAGISTER Thesis, Développement de modèles de fusion et de classification contextuelle d'images satellitaires par la théorie de l'évidence et la théorie du raisonnement plausible et paradoxal, by Nassim ABBAS, University of Science and Technology Houari Boumediene (USTHB), BP. 32, El Alia, Bab Ezzouar, 16111, Algiers, Algeria, 05 March 2009.
14. 2009 – France - Ph.D. Thesis, Prise en compte de l'incertitude dans l'expertise des risques naturels en montagne par analyse multicritères et fusion d'information, Jean-Marc Tacnet, Ecole Nationale Supérieure des Mines de Saint-Etienne, Laboratoire SITE, France, November 26th, 2009. (in French)
15. 2009 – China – Ph.D. Thesis, DSmT-based Trust Management in Open Computational Systems, J. Wang, Nanjing University of Technology and Engineering.

16. 2009 – China – Ph.D. Thesis, Information Fusion Technology Based on DSmT and Its Application in the Map Building for Robot, J. Gao, Huazhong University of Science and Technology.
17. 2009 – France - HDR Thesis, Modélisation et gestion du conflit dans la théorie des fonctions de croyance, by A. Martin, Université de Bretagne Occidentale, 23 novembre 2009. (in French)
18. 2009 – Canada - Ph.D. Thesis, Cartographie de paramètres forestiers par fusion évidentielle de données géospatiales multi-sources: Application aux peuplements forestiers en regénération et feuillus matures du sud du Québec, Brice Mora, Sherbrooke Univ., Canada, March 5th, 2009.
19. 2010 – France - Ph.D.Thesis, Extraction de connaissances et indexation de données multimédia pour la détection anticipée d'évènement indésirable, by Anas Dahabiah, Télécom Bretagne & Université Rennes 1, Brest, France, October 8th, 2010.
20. 2010 – Poland - Ph.D. Thesis, Processing of information in C2 systems with application of Dezert-Smarandache Theory, Ksawery Krenc, Military University of Technology (WAT), The Department of Electronics, Supervisor Prof. Adam Kawalec, Warsaw, Poland, 10 June 2010.
21. 2010 – Netherlands - Ph.D. Thesis, Sensing, what matters?, W. Van Norden (CAMS -- Force Vision), Delft Techn. Univ., Netherlands, 16 February, 2010.
22. 2010 – China – Ph.D. Thesis, Incomplete Information Fusion Technology and Its Application in Mobile Robot, by P. Li, Huazhong University of Science and Technology.
23. 2010 - USA - Ph.D. Thesis, Context reasoning under uncertainty based on evidential fusion networks in home-based care, by Hyun Lee, University of Texas at Arlington, USA, Aug. 2010.
24. 2011 – France - HDR Thesis, Planification de capteurs et fusion de l'information dans un système d'observation et de renseignement, by F. Dambreville, University of Bretagne Occidentale, ENSTA-Brest, France, November 16th, 2011. (in French)
25. 2011 – Romania - Ph.D. Thesis, Contributii privind monitorizarea retelelor de calculatoare, Nicu-Sebastian Nicolaescu, Academia Tehnica Militara, Bucharest, Romania, 03 November 2011. (in Romanian)
26. 2013 – Canada - Ph.D. Thesis, Optimisation d'algorithmes d'approximation de fonctions de croyance généralisées, Pascal Djiknavorian, Laval University, Québec, Canada, November 15th, 2013. (in French)
27. 2013 - Morocco - Ph.D. Thesis, Fusion d'images par la théorie de Dezert-Smarandache (DSmT) en vue d'applications de télédétection, Azeddine El Hassouny, University Ibn Zohr, Agadir, Morocco, June 22, 2013.
28. 2013 – Romania – Ph.D.Thesis, Contributions to the Development of Hybrid Force-Position Control Strategies for Mobile Robots Control, Eng. Ionel Alexandru Gal, Institute of Solid Mechanics, Romanian Academy, Bucharest, October 14, 2013.
29. 2014 - Germany - Ph.D. Thesis, Probabilistische Fahrzeugumfeldschätzung für Fahrerassistenzsysteme, Simon Steinmeyer, Fakultät für Elektrotechnik, Informationstechnik, Physik der Technischen Universität Carolo-Wilhelmina zu Braunschweig, Germany, Mai 13th, 2014 (in German).
30. 2014 – China - Ph.D. Thesis, Credal classification of uncertain data based on belief function theory, by Zhunga Liu, Northwestern Polytechnical University (NPU), Xi'an, China in co-tutelle with Telecom Bretagne, Brest France, defended at NPU in Xi'an on Nov. 24th, 2014.

Awards

- New Mexico-Arizona Book Award at the category Science & Math, for the book "DSm Super Vector Space of Refined Labels", by W. B. Vasantha Kandasamy, F. Smarandache, 16 November 2012, Albuquerque, NM, USA.
- Best paper award for the paper entitled On the Behavior of Dempster's Rule of Combination and the Foundations of Dempster-Shafer Theory, by Tchamova A., Dezert J., presented at IEEE IS'2012, Sofia, Bulgaria, Sept. 6-8, 2012.

Patent

- DSmT (Dezert-Smarandache Theory)-based image target multi-characteristic fusion recognition method, by Xinde Li, Weidong Yang, European Pattent Office, Espacenet, Bibliographic data: CN102222240 (A), 2011-10-19.

Biographies of contributors

Nassim Abbas was born in Algeria in 1981. From 2000 to 2002 he achieved his studies of fundamental sciences at the National Polytechnic School, Algiers, Algeria. He obtained his engineering degree in electronics from the Faculty of Electronics and Computer Science, University of the Sciences and Technology Houari Boumediene (USTHB), Algiers, Algeria, in 2006. He received his Magister and Ph.D. degrees in Signal and Image Processing from USTHB in 2009 and 2015, respectively. From 2011 to 2012, he was a research associate in the Systems Architecture and Multimedia division, Center for Development of Advanced Technologies, Algiers, Algeria. In 2012, he joined the Department of Telecommunication, Faculty of Electronics and Computer Science, at USTHB where he is currently a teacher-researcher since 2013. His research interests include feature generation, machine learning, evidence theory, plausible and paradoxical reasoning theory and its applications for document analysis, multimodal biometric verification, and satellite images classification and fusion.
Email : nabbas(at)usthb.dz

Youssif Al-Nashif is an Assistant Professor in the Department of Electrical and Computer Engineering at the Old Dominion University since July 2014. He received his Ph.D. in Computer Engineering from The University of Arizona in 2008. He received his B.Sc. Electrical Engineering and M.Sc. Communication and Electronics Engineering form Jordan University of Science and Technology in 1999 and 2000 respectively. His research interests include Cybersecurity, Cyber Resilience, Secure Critical Infrastructures, Cyber Trust, Cloud Security, Autonomic Computing, Data Analysis, Behavior Modeling, and Power and Performance Management.
Email : alnashif(at)ece.arizona.edu

Jin An was born in ShiJiaZhuang, China, on Feb. 8, 1981. He received BE degree from the Harbin Institute of Technology in 2003, and ME degree from the Jiangsu Automation Research Institute (JARI) in 2006 respectively. He is a senior engineer in JARI. He joined the second session of the Committee of Information Fusion Branch in Chinese Society of Aeronautics and Astronautics (CSAA) in 2014. He worked as a researcher in numerous academic and industry projects in the area of Multi-sensor Multi-target Tracking (MS-MTT) and Data Fusion.
Email : anjin308(at)163.com.

Mireille Batton-Hubert is a professor in *Modelling applied to Environmental Sciences and engineering* at the École des Mines de Saint-Étienne - institute Mines Telecoms in industrial DEcision-making, Modeling and Optimization DEMO team - Henri Fayol Department of École des Mines de Saint-Étienne. She stayed 12 years in the Information, Decision and Environmental Department, at the École Nationale Supérieure des Mines de Saint-Étienne In 2010, M. Batton-Hubert jointed a Applied mathematics laboratory to teach and to develop linked between the computation sciences, statistics tools to identify and to quantify the uncertainty propagation. She has been the head of DEMO laboratory since 2014. Her activities of research concern the development of tools for the identification, modeling, and optimization applied to living and eco-industrial processes. The scientific disciplines are the: deterministic modelling (PDE-PDO), Simulation (CFD), Optimization, and the Data mining, around applied mathematics for decision-making. These activities concern the forecasting

of impact and the risk assessment and health monitoring, the process control, and the prevision) with regard to different physical process (fluid dynamic). The applied research topics concern: aerodynamics - hydrodynamics and groundwater - impacts hydro territorial adjustments - simultaneous analysis of hazards operating on drinking water supply network - time distribution of air quality and atmospheric pollution and odour prediction - information fusion for natural risk analysis. Her research activities have been acknowledged by Scientific references for 2005 - 2014: 31 papers with 8 specific books and chapters and 2 software registration IDDN, 14 technical reports for projects management with companies.
Email : batton(at)emse.fr, Mireille.BATTON-HUBERT(at)emse.fr

Kaouthar Benameur received her PhD in Electrical Engineering from McGill University in 1997 and her MSc in Electrical Engineering in 1990 from École Polytechnique, both in Montreal, Canada. Dr. Benameur is currently working as a senior research engineer at ONERA. Her research interests include tracking systems under communication constraints, distributed information fusion and dynamic resource management, modeling and control of sensor networks.
Email : kaouthar.benameur(at)onera.fr

Erik Blasch (IEEE S'98-M'99-SM'05) is a principal scientist at the the United States Air Force Research Laboratory (AFRL) in the Information Directorate at Rome, NY, USA. From 2009-2012, he was an exchange scientist to Defence Research and Development Canada (DRDC) at Valcartier, Quebec. From 2000-2009, Dr. Blasch was the Information Fusion Evaluation Tech Lead for the AFRL Sensors Directorate - COMprehensive Performance Assessment of Sensor Exploitation (COMPASE) Center supporting design evaluations. Dr. Blasch was previously an Adjunct Electrical Engineering Professor at Wright State University and the Air Force Institute of Technology in Dayton, Ohio teaching signal processing, target tracking, and information fusion (2000-2010). Dr. Blasch currently serves as an Institute of Electrical and Electronics Engineers (IEEE) Aerospace and Electronics Systems Society (AESS) Board of Governors (BoG) member (2011-2016), the AESS International Chapters Chair, AESS representative to AIAA (2013-2016), and an AESS Distinguished lecturer (2013-2016). Dr. Blasch was a founding member of the International Society of Information Fusion (ISIF) (www.isif.org) in 1998, held various leadership roles in the Fusion conferences, a BoG member (1999-2010), and the 2007 ISIF President. He has focused on robotics, pattern recognition, targeting tracking, and information fusion research compiling 500+ scientific papers, tutorials, and book chapters. His recent book is *High-Level Information Fusion Management and Systems Design* (Artech House, 2012). Dr. Blasch received his B.S. in Mechanical Engineering from the Massachusetts Institute of Technology in 1992 and Master's Degrees in Mechanical ('94), Health Science ('95), and Industrial Engineering (Human Factors) ('95) from Georgia Tech and attended the University of Wisconsin for a MD/PhD in Neurosciences/Mech. Eng until being called to military service in 1996 to the United States Air Force. He completed an MBA ('98), MSEE ('98), MS Econ ('99), and a PhD ('99) in Electrical Engineering from Wright State University and is a graduate of Air War College ('08). He is the recipient of the IEEE Bio-Engineering Award (Russ-2008), IEEE AESS Magazine Best Paper Award (Mimno-2012), Military

Sensing Symposium Leadership in Data Fusion Award (Mignogna-2014). He is a Fellow of SPIE, Associate Fellow of AIAA, and a senior member of IEEE.
Email : erik.blasch.1(at)us.af.mil, erik.blasch(at)rl.af.mil

Nadejda Bocheva is Associate Professor at Institute of Neurobiology (former Institute of Physiology), Bulgarian Academy of Sciences. She received a BS degree in solid state physics and MS degree in biophysics from Sofia University in 1979 and a Ph. Degree in biology in 1986 from Institute of Physiology, Bulgarian Academy of Sciences. In 1979 she joined the research team "Mechanisms of perception" at the Institute of Physiology, Bulgarian Academy of Sciences. Since 1986 she worked there as an Assistant Professor, and in 2006 became Associate Professor in psychophysiology. At present she is head of Department of Sensory Neurobiology at the Institute of Neurobiology. Her research interests are in human visual information processing, spatial vision, motion perception, visual recovery of 3D shape and ageing. She is a Fulbright fellow and has a Fogarthy international collaborative Award in 2002. She is member of American Psychological Association and of the Sofia section of the Bulgarian Physiological Society.
Email : nadya(at)percept.bas.bg; nbbocheva(at)hotmail.com

Yong-mei Cheng was born in 1960, Xi'an China. She received her Bachelor degree in Xi'an University of Technology, Xi'an, China, and her Master and PhD degrees were respectively obtained in Northwestern Polytechnical University (NPU) in 1997 and 2001. She is working in School of Automation, NPU since 1994, and she is Professor in NPU. Her research work mainly focus on Signal and Information Processing, information fusion Maneuver Multiple Target Tracking, Artificial Intelligence and pattern recognition.
Email : chengym(at)nwpu.edu.cn
Web page : http://zdhxy.nwpu.edu.cn/info/1105/1889.htm

Youcef Chibani received his Magister and Doctoral Degrees in electrical engineering from the University of Science and Technology Houari Boumedienne, Algiers, Algeria. Currently, he is working as a teacher-researcher, as a full Professor since 2007. His research interests include character recognition, segmentation, feature extraction, multidirectional decomposition, image enhancement neural network, support vector machines and its application for document analysis, writer identification and signature verification. He was a Head of the Scientific Council and responsible for master and doctoral training. He co-authored many papers in international peer-reviewed journal and conferences. More recently, he is director and founding member of a new laboratory namely Laboratoire d'Ingéniérie des Systèmes Intelligents et Communicants (Communicating and Intelligent System Engineering Laboratory).
Email : ychibani(at)usthb.dz
Web page: http://www.lisic.usthb.dz

Franco Ciaramaglia was born in Gaeta, Italy, on 22nd May, 1962. He received a Doctorate in Electronic Engineering from the University of Rome (I), Italy, in 1988. In 1990, he joined "Officine Galileo" (Florence), now SELEX Electronic Systems, in the Environmental System Department, where he was an analyst of Geographical Information Systems and a Remote Sensing specialist from 1995. Following this, he moved to the Avionics Department and was

involved in many national and international projects in civil avionics, mission and airborne surveillance systems. Currently, he leads the Innovation and Research team dealing with Remotely Piloted Aircraft (RPA). His main research interests are now in the area of Data/Sensors Fusion, Knowledge Management and System Autonomy.
Email : franco.ciaramaglia(at)selex-es.com

Paulo C. G. Costa is Associate Professor in the Department of Systems Engineering and Operations Research and Director of the Radio and Radar Lab, both at George Mason University. He received his Ph.D. in Information Technology from GMU and has been an active researcher and lecturer in the areas of probabilistic ontologies, decision support systems, cyber security, multi-sensor data fusion, modeling and simulation, and probabilistic representation and reasoning. Dr. Costa has an extensive academic service, including acting as chair of conferences such as the 18th Information Fusion and working groups, such as ISIF's Evaluation of Uncertainty Reasoning Technologies working group. He has been an active member of engineering societies such as IEEE, ISIF and the International Council of Systems Engineering. He is also Research Director - C2 International Activities of the GMU's Center of Excellence in C4I, an IEEE senior member, and holds courtesy affiliations with the University of Brasilia and the Instituto Tecnológico da Aeronáutica (ITA-Brazil).
Email : pcosta(at)c4i.gmu.edu
Web page : http://mason.gmu.edu/~pcosta/

Jean Dezert was born in l'Hay les Roses, France, on August 1962. He received the EE degree in 1985, and his Ph. D. from the University Paris XI, Orsay, in 1990 in Automatic Control and Signal Processing. During 1986-1990 he was with the Syst. Dept. at the French Aerospace Lab (ONERA) and did research in multisensor multitarget tracking (MS-MTT). During 1991-1992, he visited the Dept. of ESE, UConn, Storrs, USA as an ESA Postdoctoral Research Fellow under supervision of Prof. Bar-Shalom. During 1992-1993 he was teaching assistant in EE Dept, University of Orléans, France. Since 1993, he is senior research scientist in the Information Modeling and Processing Dept. at the French Aerospace Lab, Palaiseau, France. His current research interests include estimation theory, and information fusion (IF) and plausible reasoning with applications to MS-MTT, object classification, pattern recognition, defense and security, robotics and risk assessment. Jean Dezert has been involved within ISIF (www.isif.org) since its beginning and has been the Local Arrangements Co-Organizer of the first Fusion Conference in Europe in 2000. He has served as vice-president of ISIF in 2004 and has been involved in the Technical Program Committees of Fusion 2001-2015 Conf., and in several sessions and panel discussions on reasoning under uncertainty and data fusion. Jean Dezert has published four books and more than a hundred of papers in conferences and journals on tracking and information fusion. He has given seminars and workshops in the IF and tracking in North America, Europe, Australia and China and he is reviewer for several international journals and Associate Editor of ISIF Journal of Advances in Information Fusion. Jean Dezert is the co-founder with Florentin Smarandache of DSmT (Dezert-Smarandache Theory) of information fusion based on belief functions.
Email : Jean.dezert (at) onera.fr, jdezert (at)gmail.com
Web page : http://www.onera.fr/staff/jean-dezert

Pascal Djiknavorian received a Ph.D. in information fusion from Laval University. He has also completed a M.Sc. in electrical engineering on information fusion applied to ESM reports under STANAG 1241. He holds a Bachelor of Engineering in computer engineering and a certificate in business administration from Laval University. Since January 2014, he is Senior Systems Architect, Product Research and Development in Business Intelligence at Allagma Technologies. His Research and Development projects there leads him to work on the design, planning and development of new technologies which requires notions of cryptography, secure communications, database, taxation, economics, medical sciences, patent development and intellectual property along with team & project management. He also occasionally consults on Research and Development projects. His Ph.D. in information fusion was a Study of Approximation Algorithms of Generalized Belief Functions, under the direction of Professors Dominic Grenier and Pierre Valin. During his Ph.D. he had the opportunity to collaborate with Prof. Arnaud Martin from the Superior National School of Advanced Techniques of Brittany. He has a dozen publications in books, journals, and conference's proceedings. His research interests include evidential theory, approximation algorithms, optimization methods, metrics development, vehicular intelligence, and communication systems among other things. Among his specialties we find intelligent systems, Mathematical Theory of Evidence, Decision Support Systems, Reasoning under uncertainty, and Pattern Recognition.
Email : pascal.djiknavorian(at)gmail.com

Valentina Dragos is a research scientist with ONERA, The French Aerospace Lab. Her research is in the field of semantic technologies applied to high-level information fusion and includes interests in culturally motivated aspects of creating and using semantically enabled procedures for soft-data analysis and decision making. She received a Ph.D. in Computer Science from the Réné Descartes University (Paris), in 2007.
Email : valentia.dragos(at)onera.fr

Alfonso Farina received the doctor degree in electronic engineering from the University of Rome (I), Italy, in 1973. In 1974, he joined Selenia, now SELEX Electronic Systems, where he has been a manager since May 1988. From 1979 to 1985, he was also a professor of radar techniques with the University of Naples; in 1985, he was appointed associate professor. He is the author of more than 500 peer-reviewed technical publications and the author of books and monographs: *Radar Data Processing* (Vol. 1 and 2) (translated in Russian and Chinese; United Kingdom: Researches Studies Press, and New York: 6 Wiley, 1985–1986); *Optimized Radar Processors* (on behalf of IEE, London, United Kingdom: Peter Peregrinus, 1987); and *Antenna Based Signal Processing Techniques for Radar Systems*, 1992. He wrote the chapter "ECCM Techniques" in the *Radar Handbook* (2nd ed., 1990, and 3rd ed., 2008), edited by Dr. M. I. Skolnik (Naval Research Laboratory, USA). In 1987, he received the Radar Systems Panel Award of IEEE Aerospace and Electronic Systems Society (AESS) for development of radar data processing techniques. He is the Italian representative of the International Radar Systems Panel of the IEEE AESS. He is vice president for industry in the board of governors of IEEE AESS. He has been the Italian industrial representative (panel member at large) at the Sensor and Electronic Technology (SET) of Research Technology Organization (RTO) of North Atlantic Treaty Organization (NATO). He has been on the board of directors of the International Society for Information Fusion (ISIF). He has been

the executive chair of the International Conference on Information Fusion (Fusion) 2006, Florence, Italy, July 10–13, 2006. He has been nominated fellow of IEEE with the following citation: "For development and application of adaptive signal processing methods for radar systems." Recently, he has been nominated international fellow of the Royal Academy of Engineering, United Kingdom; this fellowship was presented to him by HRH Prince Philip, the Duke of Edinburgh. In September 7, 2006, he received the Annual European Group Technical Achievement Award 2006 by the European Association for Signal, Speech and Image Processing (EURASIP), with the citation: "For development and application of adaptive signal processing technique in practical radar systems." He has been the general chairman of the IEEE Radar Conference 2008, Rome, May 26–30, 2008. He is a fellow of the Institution of Engineering and Technology (IET), United Kingdom. He has been recently nominated fellow of EURASIP; the citation reads: "For contributions to radar system design, signal, data and image processing, data fusion and particularly for the development of innovative algorithms for deployment into practical radar systems." He is also the recipient of the 2010 IEEE Dennis J. Picard Gold Medal for Radar Technologies and Applications with the following citation: "For continuous, innovative, theoretical and practical contributions to radar systems and adaptive signal processing techniques." He is the recipient of the IET Technical Achievement Awards in 2014. He retired from Selex ES in 2014. Today he is private teaching and consulting in radar.
Email : alfonso.farina.it(at)ieee.org

Giovanni Golino received the doctor degree in electronic engineering cum laude from the University of Rome (II) in 1998. In 1999 he joined former Alenia Marconi Systems, now Selex Electronic System, as radar system analyst. He has worked as analyst and designer of radar signal processing algorithms up to 2011. He was co-recipient in 2008 of the Innovation Award from former Selex Sistemi Integrati, now Selex Electronic System, for "having determined an effective business enhancement to Selex Sistemi Integrati efficiently using the high technological potential of the Company" in the project LYRA 10. He has been working as integrated system analyst and designer since 2012 and has been involved in the design of algorithms concerning multi target tracking and classification and multi sensor data fusion. His field of interest also includes machine learning, data mining and decision making under uncertainty.
Email : giovanni.golino(at)selex-es.com

Antonio Graziano is currently Head of Systems Analysis within the Land and Naval Division of Selex ES (Italy). He holds a Degree cum Laude in Electronic Engineering and has over twenty five years of experience in air defense, surveillance, command and control and homeland protection systems. He has been team leader of several national and international study projects on data fusion and network centric operations. He is co-author of more than 50 peer-reviewed technical publications.
Email : antonio.graziano(at)selex-es.com

Dominic Grenier received the M.Sc. and Ph.D. degrees in electrical engineering in 1985 and 1989, respectively, from the Université Laval, Québec City, Canada. From 1989 to 1990, he was a Postdoctoral Fellow in the radar division of the Defense Research Establishment in Ottawa (DREO), Canada. In 1990, he joined the Department of Electrical Engineering at

Université Laval where he is currently a Full Professor since 2000. Recognized by the undergraduate students in electrical and computer engineering at Université Laval as the electromagnetism and RF specialist4. His excellence in teaching has resulted in "Best Teacher Award" directly from students' associations many times, and the Excellence-teaching award from Université Laval in 2010; finally he received a special teaching recognition award in 2012 from "Ordre des ingénieurs du Québec" (OIQ). He obtained in 2009 one special teaching chair from the Quebec Minister for education. His research interests include inverse synthetic aperture radar imaging (ISAR), array signal processing for high resolution estimators, reflectometry probes design, and information fusion (adaptive combination rules in evidence theory and DSm theory, hybrid reasoning frame, DSmT apply to Interrogated Friend or Foe sensor informations). Prof. Grenier has close to 40 publications in refereed journals and 75 more in conference proceedings. In addition, more than 40 graduate students completed their thesis under his direction since 1992.
Email : dominic.grenier(at)gel.ulaval.ca

Bilal Hadjadji received the M.S. degree in 2012 from the Faculty of Electronics and Computer Sciences, University (USTHB), Algiers, Algeria,. Currently he is a Ph.D student at the same faculty. His research interests include machine learning, multiple classifiers system, image processing, biometrics and handwritten recognition.
Email : bhadjaji(at)usthb.dz

Chongzhao Han was born in 1943, in Shaanxi, PR China. He is now a Professor in School of Electronic and Information Engineering, Xi'an Jiaotong University. His research fields include information fusion, automatic control and nonlinear system analysis. He is an ISIF member. He was the vice president of the TPC of the 1st -6th Chinese Conference on Information Fusion held in 2009-2014.
Email : czhan(at)mail.xjtu.edu.cn

Deqiang Han was born in Xi'an, Shaanxi, China, on March 03, 1980. He received the Bachelor Degree of Communication and Control engineering from Xi'an Jiaotong University in 2001, the Master Degree of Control Science and Engineering from Xi'an Jiaotong University in 2004. He got his Ph.D Degree of Control Science and Engineering in 2008 from Xi'an Jiaotong University. He is currently an Associate Professor at Xi'an Jiaotong University, China. He is an IEEE member and ISIF member. His research fields include evidence theory, information fusion and pattern classification. He was the IPC member of the 1st -6th Chinese Conference on Information Fusion held in 2009-2014.
Email : deqhan(at)mail.xjtu.edu.cn, dqhan(at)gmail.com

Salim Hariri is a Professor in the Department of Electrical and Computer Engineering at The University of Arizona. He received his Ph.D. in computer engineering from University of Southern California in 1986, and an MSc from The Ohio State University in 1982. He is the UA site director of NSF Center for Cloud and Autonomic Computing and he is the Editor-In-Chief for the CLUSTER COMPUTING JOURNAL (Springer, http://clus.edmgr.com) that presents research techniques and results in the area of high speed networks, parallel and distributed computing, software tools, and network-centric applications. He is the Founder of the IEEE/ACM International Symposium on High Performance Distributed Computing

(HPDC) and the co-founder of the IEEE/ACM International Conference on Cloud and Autonomic Computing. He is co-author/editor of four books on Autonomic computing, parallel and distributed computing: Autonomic Computing,: Concepts, Infrastructure, and Applications (CRC Press, 2007), Tools and Environments for Parallel and Distributed Computing (Wiley, 2004), Virtual Computing: Concept, Design and Evaluation (Kluwer, 2001), and Active Middleware Services (Kluwer, 2000). His research interests include autonomic cyber security, big data analytics, resilient cloud services, critical infrastructure protections, and autonomic programming, and resilient Dynamic Data Driven Application Systems (rDDDAS).
Email : hariri(at)ece.arizona.edu
Web page : nsfcac.arizona.edu

Xinhan Huang was born in Hubei province, China, on August 21, 1946. He graduated from Huazhong University of Science and Technology (HUST), Wuhan, China in 1969. He is Faculty member of the Department of Control Science and Engineering of HUST. He joined the Robotics Institute of Carnegie-Mellon University at Pittsburgh, USA as a visiting scholar from 1985 to 1986 and the Systems Engineering Division of Wales University at Cardiff, UK as a senior visiting scholar in 1996. He is currently a Professor and Head of the Intelligent Control and Robotics Laboratory of HUST. The Chinese National Natural Science Foundation and the National High Technique Research Developing Plan award Professor Huang for his research projects from 1988 to 2008. His research interests are the areas of robotics, sensing techniques, information fusion and intelligent control. He has more than 300 research publications to his credit. Professor Huang is a senior member of Chinese Association for Artificial Intelligence (CAAI) and Chairman of the Speciality Committee of Intelligent Robotics.
E-mail : xhhuang(at)mail.hust.edu.cn

Audun Jøsang joined Oslo University in 2008. Prior to that he was Associate Professor at QUT, research leader of Security at DSTC in Brisbane, worked in the telecommunications industry for Alcatel in Belgium and for Telenor in Norway. He was also Associate Professor at the Norwegian University of Science and Technology (NTNU). He has a Master's in Information Security from Royal Holloway College, University of London, and a PhD from NTNU in Norway. He is a CISSP and CISM with broad knowledge and experience in information security.
Email : josang(at)mn.uio.no
Web page : http://www.mn.uio.no/ifi/english/people/aca/josang/

Anne-Laure Jousselme is with the Centre for Maritime Research and Experimentation (CMRE) in La Spezia (Italy) since June 2014. She was previously with the Defence Research and Development Canada (DRDC) in Valcartier in the Command, Control and Intelligence (C2I) section. She received her Ph.D in 1997 jointly from the Institut National Polytechnique de Grenoble (France) and the Electrical Engineering and Computer Engineering Department of Laval University in Quebec City (Canada). She is a member of the boards of directors of the International Society of Information Fusion (ISIF) and of the Belief Functions and Applications Society (BFAS) since 2012. She is a member of the steering committee of the BELIEF conference and an active member of the ISIF working group on

Evaluation Techniques of Uncertainty Representation (ETUR). She is member of the program committees for the Information Fusion conference (FUSION), the Canadian Tracking and Fusion Group workshop (chair of the 2014 edition), the International Conference on Belief Functions (BELIEF), the IEEE Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA) and the International Conference on Semantic Technologies for Intelligence, Defense, and Security (STIDS). She was a member of the organizing committee for the Fusion conference held in Quebec City in 2007, for which she served as tutorial chair and co-organized a special session on *Information Fusion and Pattern Recognition*. She has been adjunct professor at Laval University (Quebec, Canada) during 10 years where she supervised master and PhD theses in the field of information fusion. Her current research interests include high-level information fusion, reasoning under uncertainty, maritime anomaly detection, target recognition and identification.
Email : Anne-Laure.Jousselme(at)cmre.nato.int

Adam Kawalec was born in Poland in 1949. He received the M.Sc. degree in solid state electronics and the Ph.D. from Faculty of Technical Physics and D.Sc. degrees in electronics, acoustoelectronics from Faculty of Electronics, Military University of Technology, Warsaw, Poland in 1974, 1980, 2002, respectively. Professor title he received from the President of Poland in 2010. In 1974, he joined the Military University of Technology, Faculty of Technical Physics, Warsaw, Poland, where he was involved in surface acoustics waves (SAW) convolvers research. Since 1979 he has tested SAW dispersive delay lines applied in radar pulse compression systems and SAW sensors and its signal processing aspects. In these studies a number of valuable results was obtained, in particular with the range of dispersion of broadband lines using the higher harmonics. Since 2002 he is involved in research in the field of signal processing, in particular, Space-Time Adaptive Signal Processing for sea surveillance radars, the effectiveness of the position estimation algorithms of microwave emission sources in the passive location system technical observation of electromagnetic radiation sources and concept, applied methods (an application of Dezert-Smarandache Theory) and techniques of Information Fusion with application in Command and Control systems. In the years 1974-2002 he held the position of teaching and the organization of senior engineer to assistant professor and head of the laboratory of the Institute of Technical Physics to the head of the Department of Planning and Coordination of Research, the Military University of Technology. In 2003-2012 professor Kawalec was the head of Institute of Radar Technology and in 2013-2014 he held the position director of Center for Technology Transfer in Military University of Technology, Warsaw, Poland. He is currently the Full Professor at the Faculty of Electronics Military University of Technology, Warsaw, Poland, where he teaches courses in acoustoelectronics and the numerical methods, theory of vibration and field and waves electromagnetic. He is the author and co-author of more than 260 scientific papers published in international and national journals and conference proceedings and co-inventor of five patents. Prof. Kawalec is a member of many scientific committees of major international conferences. He is Senior Member of IEEE organization.
Email : Adam.Kawalec(at)wat.edu.pl

Pavlina Konstantinova received her M.Sc. degree in Computer Science from Technical University of Sofia in 1967 and PhD degree in Computer-aided design of nonlinear

electronic circuits 1987. Her research interests include sensor data processing, data association, mathematical modelling, image processing and parallel algorithms. She is now Associate Professor at the Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Sofia, Bulgaria.
E-mail : pavlina.konstantinova(at)gmail.com, pavlina(at)bas.bg
Web page : http://mmsip.bas.bg/staff/pavlina.html

Ksawery Krenc is Assistant Professor at R&D Marine Technology Centre (CTM). He received M.Sc. in Automation and Robotics, Faculty of Electronics Telecommunication and Informatics from The Technical University of Gdansk in 2001 and Ph.D. in Telecommunication from Military University of Technology (MUT) in 2010. Since 2008 he has been active in the field of Information Fusion, including publishing solutions (methods, techniques, and algorithms) mostly for the purpose of Command and Control systems. Since 2011 has been participating as a member in NATO STO (IST) task groups, as well as NIAG groups. He has served as a reviewer of papers for Elsevier's Information Fusion journal, and since 2014 as a member of technical committee for Fusion conference. His current research interests focus on Information fusion for the purpose of Command and Control systems as well as Computer-Aided Detection/Computer-Aided Classification (CAD/CAC) solutions for mine warfare, with a particular emphasis on applying evidential theories like Dempster-Shafer Theory or Dezert-Smarandache Theory. As a member of STO groups he is also interested in application of Battle Management Language (BML) in multi-level fusion systems, concentrating on uncertainty management problems.
Email : ksawery.krenc(at)ctm.gdynia.pl,ksawery.krenc(at)ctm.gdynia.pl

Valeri Kroumov is a Professor at the Okayama University of Science, Faculty of Engineering, Department of Electrical Engineering, Okayama, Japan. He is Director of the Research Institute of Technology within this university. He published many papers on robotics and presented papers to various international conferences on mechatronics. He is the editor-in-chief of the Bulletin of the Research Institute of Technology (in Japanese language).
Email : val(at)ee.ous.ac.jp
Web page : http://shiwasu.ee.ous.ac.jp/val-e.htm

Kathryn Blackmond Laskey is Professor of Systems Engineering and Operations Research and Associate Director of the Center of Excellence in Command, Control, Communications, Computing and Intelligence at George Mason University, Fairfax, Virginia, where she teaches undergraduate and graduate courses in systems engineering, artificial intelligence, and statistics. She performs research on decision theoretically based knowledge representation and inference, decision and inference support systems, learning in high-dimensional parameter spaces, and multi-source information fusion. Professor Laskey received a joint PhD in statistics and public affairs from Carnegie Mellon University, an MS in mathematics from the University of Michigan, and a BS in mathematics from the University of Pittsburgh.
Email : klaskey(at)c4i.gmu.edu
Web page : http://seor.gmu.edu/~klaskey/

Xinde Li got this Ph.D in Control Theory and Control Engineering, from Department of Control Science and Engineering, Huazhong University of Science and Technology (HUST), Wuhan, China, in 2007. Afterwards, he joined School of Automation, Southeast University, Nanjing, Jiangsu, China, where he is currently Associate Professor and Ph.D. students Supervisor. His research interests include information fusion, object recognition, computer vision, intelligent robot and human-robot interaction, etc. He has published more than 60 SCI and EI papers and holds 9 9 national patents. He was granted a «Talent of Qing Lan Project » award of Jiangsu province, and a « Six Major Top-talent Plan » award of Jiangsu province, China.
Email : xindeli(at)seu.edu.cn

Yanli Li was born in 1969 in China. She graduated from the department of applied mathematics in Nanjing University of Science and Technology (NUST) in 1991. She received her master's degree in department of systems engineering at China Ship Research and Development Academy (CSRDA), in 1994. She has long been engaged in the research of target identification. She is an author of more than 20 papers and has obtained one patent for target identification. Currently, she is a researcher at Jiangsu Automation Research Institute (JARI). She has extensively worked in areas of information fusion, target identification and C4ISR.
E-mail : liyanli_jari(at)163.com

Zhicheng Liu was born in Yangzhou, China, in 1984. He received his bachelor's degree from the Department of Mathematics at Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China in 2006, and his master's degree in department of systems engineering at Jiangsu Automation Research Institute (JARI) in 2009. After graduation, he joined the team of information fusion in JARI. He has been an author of more than 10 papers and three of his papers have been retrieved by EI. His research interests include Information Fusion, Target Identification Based on Radiation and Target Synthetic Identification.
E-mail : zhichengliu2002(at)163.com

Zhunga Liu was born in Luoyang, China, 1984. He received the Bachelor and Master degree from Northwestern Polytechnical university (NPU), Xi'an, China in 2007 and 2010. His Ph.D degree was obtained from NPU and Telecom Bretagne in France in 2014. He is now Associate Professor in School of Automation, NPU. His research interest mainly focuses on belief function theory and its application in pattern recognition especially for data classification.
Email : liuzhunga(at)gmail.com, liuzhunga(at)nwpu.edu.cn

Arnaud Martin is full professor at University of Rennes 1 in the team DRUID of IRISA laboratory. He received a HDR (French ability to supervised research) in computer sciences (2009), a PhD degree in Signal Processing (2001), and Master in Probability (1998). Dr. Arnaud Martin worked on speech recognition during three years (1998-2001) at France Telecom R&D, Lannion, France. He worked in the department of statistic and data mining (STID) of the IUT of Vannes, France, as temporary assistant professor during two years (2001-2003). In 2003, he joined the laboratory E3I2 at the ENSTA Bretagne, Brest, France,

as a teacher and researcher until 2010. In 2010, he joined the laboratory IRISA at the university of Rennes 1 as full professor and create the team DRUID in 2012. Pr. Arnaud Martin is the head of the computer science department at the IUT of Lannion and teaches mathematics, data fusion, data mining, signal processing and computer sciences. His research interests are mainly related to the belief functions for the classification and include data fusion, data mining. He is author of numerous papers and invited talks. He supervised numerous PhD students. Pr. Arnaud Martin is the founder in 2010 of the Belief Functions and Applications Society (BFAS) (www.bfasociety.org). From 2010 until 2012 he was the president of BFAS, and since 2012 he is the treasurer.
Email : Arnaud.Martin(at) univ-rennes1.fr
Web page : http://www.arnaud.martin.free.fr/

Valter Mellano was born in Turin, Italy, on April 16 1957. He received the Physics degree in 1981 from the University of Turin. He joined to Aeritalia, an aeronautical company, in 1983, as a SW Engineer. He works in the same company, now named Finmeccanica Selex ES. He spent the first years in Research & Development about Artificial Intelligence, then he was busy in SW development projects. Therefore he got a long experience in Avionics System Engineering, and took part into different international programs. Since 2009 he has been involved in Unmanned Aerial System (UAS) development. Now he is the technical responsible for the development and integration of Research projects into the new UAS products.
Email : walter.mellano(at)selex-es.com

Grégoire Mercier, IEEE M'99, SM'07, was born in France in 1971. He received the Engineer Degree from the Télécom SudParis, Évry, France in 1993, his Ph.D. degree from the University of Rennes I, Rennes, France in 1999 and his "Habilitation à Diriger des Recherches" from the University of Rennes I in 2007. Since 1999, he has been with Télécom Bretagne, where he has been an Associate Professor in the Image and Information Processing department (ITI). He is Professor since 2010. His research interests are in remote sensing image compression and segmentation, especially in hyperspectral and Synthetic Aperture Radar images. At present time, his research is dedicated to change detection from image time series and to heterogeneous data processing. He was a visiting researcher at DITEN (University of Genoa, Italy) from March to May 2006 where he developed change detection technique for heterogeneous data. He was also a visiting researcher at CNES (France) from April to June 2007 to take part of the Orfeo Toolbox development. Since 2012, he is external collaborator of the AYIN research group at INRIA-SAM. Prof Mercier is has been an Associate Editor for the IEEE Geoscience and Remote Sensing Letters (GRSL) from 2007 to 2014. He has been President of the French Chapter of IEEE Geoscience and Remote Sensing Society from 2010 to 2013. He is now Associate Editor for the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS). He serves as Head of the Technical Commitee of the IEEE MultiTemp'2015 workshop.
Email : gregoire.mercier(at)telecom-bretagne.eu
Web page : http://www.gregoire-mercier.fr/

Christophe Osswald was born in Orléans, France, in 1975. He received a Ph.D. degree in Mathematics and Computer Science (2003) from the EHESS -- École des Hautes Études en Sciences Sociales -- after three years of research on classification, similarity analysis and hypergraphs at the laboratory IASC of the ENST Bretagne. He is an engineer graduated from the École Polytechnique (1997) and the ENST Bretagne (1999), and he received the Master degree in Mathematics and Computer Science applied to Social Sciences. He is a teacher and researcher at the ENSTA Bretagne in the laboratory LabSTICC: UMR6285, Brest, France. His research interests are related to data fusion, mostly through belief functions, and also to classification, graphs and hypergraphs theory, optimization problems, algorithmic complexity and decision support.
Email: Christophe.Osswald(at)ensta-bretagne.fr

Jin-Dong Pan received his master degree from Southeast University, Nanjing, China in 2014. His research interest covers information fusion and image processing.
Email : panjindong1989(at)163.com

Quan Pan was born in Shanghai, China 1961. He received the B.E. degree from Huazhong Institute of Technology in 1991, and the M.E. and Ph.D. degrees from the Automation School at Northwestern Polytechnical University (NPU) in 1991 and 1997, respectively. He is a Professor of the Automation School, and the Director of Research Institute of Control & Information of NPU from 1998. He is the Duty Dean of Graduate School of NPU from 1996 to 2002, the Duty Dean of Management School of NPU from 2002 to 2004, and the Director of Office of Development and Planning of NPU from 2004 to 2009. Now he is the Dean of the Automation School of NPU from 2009. His research interests include information fusion, hybrid system estimation theory, multi-scale estimation theory, target tracking and image processing. He is a Member of IEEE, a Member of the International Society of Information Fusion (ISIF), a Board Member of the Chinese Association of Automation, and a Member of Chinese Association of Aeronautics and Astronautics. He has published 6 books, and almost 30 international journal papers, including IEEE TAC, Automatica and IEEE TSP. He obtained the 6th Chinese National Youth Award for Outstanding Contribution to Science and Technology in 1998 and the Chinese National New Century Excellent Professional Talent in 2000.
Email : quanpan(at)nwpu.edu.cn
Web page : http://zdhxy.nwpu.edu.cn/info/1105/4406.htm

Benjamin Pannetier was born in Paris on November 30th, 1979. He received his B.Sc. in math from the University of Marne la Vallée and his Ph.D. in automatic control and signal processing from the University of Grenoble in 2006. Since 2005, he is a research engineer at the French Aerospace Lab (ONERA), Palaiseau, France. His research interests include target tracking, detection/estimation theory and data fusion for battlefield surveillance systems for the French army. He has been working on multiple target tarcking with GMTI data by taking into account geographical information context. More recently he has lead new research activity in data fusion with abandoned ground sensor network for intelligence applications.
Email : Benjamin.Pannetier(at)onera.fr

Florentin Smarandache is a full professor of mathematics at the University of New Mexico, in USA. He is a scientist, writer and cyber-artist. He did post-doctoral researches at Okayama University of Science (Japan) between 12 December 2013 - 12 January 2014; at Guangdong University of Technology (Guangzhou, China), 19 May - 14 August 2012; at ENSIETA (National Superior School of Engineers and Study of Armament), Brest, France, 15 May - 22 July 2010; and for two months, June-July 2009, at Air Force Research Laboratory in Rome, NY, USA (under State University of New York Institute of Technology). Graduated from the Department of Mathematics and Computer Science at the University of Craiova in 1979 first of his class graduates, earned a Ph. D. in Mathematics from the State University Moldova at Kishinev in 1997, and continued postdoctoral studies at various American Universities such as University of Texas at Austin, University of Phoenix, etc. after emigration. Since 2002, together with Dr. Jean Dezert from the French Aerospace Lab, worked in information fusion and generalized the Dempster-Shafer Theory to a new theory of plausible and paradoxist fusion (Dezert-Smarandache Theory): http://fs.gallup.unm.edu/DSmT.htm. In 2004 he designed an algorithm for the Unification of Fusion Theories and rules (UFT) used in bioinformatics, robotics, military. He received the 2011 Romanian Academy "Traian Vuia" Award for Technical Science (the highest in the country); Doctor Honoris Causa of Academia DacoRomana from Bucharest - 2011, and Doctor Honoris Causa of Beijing Jiaotong University (one of the highest technical universities of China) - 2011; the 2012 New Mexico - Arizona Book Award & 2011 New Mexico Book Award at the category Science & Math (for Algebraic Structures, together with Dr. W. B. Vasantha Kandasamy) on 18 November 2011 in Albuquerque; also, the Gold Medal from the Telesio-Galilei Academy of Science from England in 2010 at the University of Pecs - Hungary (for the Smarandache Hypothesis in physics, and for the Neutrosophic Logic), and the Outstanding Professional Service and Scholarship from The University of New Mexico - Gallup (2009, 2005, 2001).
Email : smarand(at)unm.edu, fsmarandache(at)gmail.com
Web page : http://fs.gallup.unm.edu/FlorentinSmarandache.htm

Jean-Marc Tacnet (Irstea - Snow avalanche engineering and Torrent Control research unit, Grenoble, France) is a senior researcher who received two Master Engineering Degrees (ENGEES, ENGREF), a M.Sc. degree in Applied Computer Science (Grenoble University) and a PhD thesis (ENSMSE) in Environmental Engineering Sciences. His activities focus on integrated risk management and expert assessment processes applied to natural phenomena in mountains. His researches deal with decision support systems for hazard, risk assessment but also risk reduction measures effectiveness analysis in a context of imperfect information and uncertainty. His disciplinary fields are multi-criteria decision-making, information fusion, artificial intelligence (fuzzy sets, possibility, belief function theories), uncertainty analysis, software engineering.
Email : jean-marc.tacnet(at)irstea.fr

Albena Tchamova is an Associate Professor at the Department of Mathematical Methods for Sensor Information Processing, Institute of Information and Communication Technologies, Bulgarian Academy of Sciences. She received her MSc in Microelectronics and PhD in Radiolocation and Navigation from the Technical University of Sofia, Bulgaria in 1978 and 1998, respectively. Her main research interests include information fusion,

Dezert-Smarandache Theory of plausible and paradoxical reasoning, decision making under uncertainty, conflicts and paradoxes. She is a member of International Society of Information Fusion (ISIF), Association of European Operational Research Societies on Fuzzy Sets (EUROFUSE).
Email : tchamova(at)bas.bg, albena.p.tchamova(at)gmail.com
Web page : http://mmsip.bas.bg/staff/albena.html

Pierre Valin : After graduating from Harvard (PhD, 1980) under Nobel Laureate Sheldon Lee Glashow, Dr. Pierre Valin taught and conducted research in theoretical physics at various universities (Moncton, McGill, Montreal) for over 12 years. He then spent over 10 years at Lockheed Martin Canada as a Senior Member of Engineering R&D to conduct research on tactical fusion systems for Canadian Forces platforms. In 2004 he joined DRDC Valcartier, where he has been Group Leader for 'Distributed Resource Management and Fusion of Information' and most recently for ``Future C2 Concepts and Structures'' till his retirement in October 2013. From 2009-2013, Dr. Valin has concurrently occupied the position of Thrust Leader for Air Command (a programmatic position). During his stay at DRDC, he has also been an adjunct professor at Université Laval and at McMaster University for the co-supervision of graduate students. His main interests were in Multi-Sensor Data Fusion requirements and design, Command and Control decision aids, SAR image processing and imagery classifiers, information reasoning frameworks, situation and threat assessment, distributed information fusion and dynamic resource management, amongst others. Dr. Valin has been involved in information fusion from its early stages in mid`90s and was President of the International Society of Information Fusion (ISIF) in 2006, which holds annual conferences. Dr. Valin has been pretty involved in the organizing committee with the 2001 ISIF conference in Montreal and in 2007 conference organized by DRDC Valcartier in Quebec City. Dr. Valin passed away on 15 April 2014.

Pei Wang is an Associate Professor at Temple University. He received his Ph.D. in Computer Science and Cognitive Science from Indiana University at Bloomington, and his M.S. and B.S. in Computer Science, both from Peking University. His research interests include artificial intelligence and cognitive science, especially on general theory of intelligence, formal models of rationality, reasoning under uncertainty, learning and adaptation, knowledge representation, and real-time decision making. He is the Chief Executive Editor of the Journal of Artificial General Intelligence.
Email : pei.wang(at)temple.edu
Web page : http://www.cis.temple.edu/~wangp/

Yi Yang was born in 1980, in Xi'an, Shaanxi, China. She got her Ph.D Degree of Control Science and Engineering in 2010 from Xi'an Jiaotong University. She is currently a lecturer at School of Aerospace, Xi'an Jiaotong University. Her main research interests are information fusion, remote sensing, uncertainty reasoning and image processing. She is an ISIF member.
Email : jiafeiyy(at)mail.xjtu.edu.cn

Xiaohong Yu was born in Nantong, China, in 1988. She received her bachelor's degree from the Department of Mathematics at Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China in 2010, and her master's degree in department of systems engineering at Jiangsu Automation Research Institute (JARI) in 2013. She is currently an assistant engineer of the team of information fusion in JARI. Her research interests include Information Fusion, Target Identification, and Data Mining.
E-mail : yuxiaohong098(at)163.com

Qingjun Zhou got his bachelor's degree in the major of Automation from Northwestern Polytechnical University (NPU), Xi'an, China, in 1991, and his master's degree in Electronics and Information Technology from Xi'an Jiao Tong University (XJTU), Xi'an, China, in 2001. He has worked in Jiangsu Automation Research Institute (JARI) since 1991. During 2001-2007, he was a Deputy Director of the C3I Department, JARI. He was appointed Director of the Department of Military Integrated Electronic Information System in 2007. He is also a graduate instructor in China Ship Research and Development Academy (CSRDA). His major research area is Theory & Technology of C4ISR System.
E-mail : chowqj(at)tom.com

The fourth volume on Advances and Applications of Dezert-Smarandache Theory (DSmT) for information fusion collects theoretical and applied contributions of researchers working in different fields of applications and in mathematics. The contributions (see List of Articles published in this book, at the end of the volume) have been published or presented after disseminating the third volume (2009, http://fs.gallup.unm.edu/DSmT-book3.pdf) in international conferences, seminars, workshops and journals.

First Part of this book presents the theoretical advancement of DSmT, dealing with Belief functions, conditioning and deconditioning, Analytic Hierarchy Process, Decision Making, Multi-Criteria, evidence theory, combination rule, evidence distance, conflicting belief, sources of evidences with different importance and reliabilities, importance of sources, pignistic probability transformation, Qualitative reasoning under uncertainty, Imprecise belief structures, 2-Tuple linguistic label, Electre Tri Method, hierarchical proportional redistribution, basic belief assignment, subjective probability measure, Smarandache codification, neutrosophic logic, Evidence theory, outranking methods, Dempster-Shafer Theory, Bayes fusion rule, frequentist probability, mean square error, controlling factor, optimal assignment solution, data association, Transferable Belief Model, and others.

More applications of DSmT have emerged in the past years since the apparition of the third book of DSmT 2009. Subsequently, the second part of this volume is about applications of DSmT in correlation with Electronic Support Measures, belief function, sensor networks, Ground Moving Target and Multiple target tracking, Vehicle-Born Improvised Explosive Device, Belief Interacting Multiple Model filter, seismic and acoustic sensor, Support Vector Machines, Alarm classification, ability of human visual system, Uncertainty Representation and Reasoning Evaluation Framework, Threat Assessment, Handwritten Signature Verification, Automatic Aircraft Recognition, Dynamic Data-Driven Application System, adjustment of secure communication trust analysis, and so on.

Finally, the third part presents a List of References related with DSmT published or presented along the years since its inception in 2004, chronologically ordered.

ISSN 1938-5544
9 771938 554002

ISBN 978-1-59973-324-1
9 781599 733241